Rational stochastic languages

François Denis, Yann Esposito

Rational stochastic languages

François Denis and Yann Esposito

LIF-CMI, UMR 6166
39, rue F. Joliot Curie
13453 Marseille Cedex 13 FRANCE
fdenis,esposito@cmi.univ-mrs.fr

Abstract. The goal of the present paper is to provide a systematic and comprehensive study of *rational stochastic languages* over a semiring $K \in \{\mathbb{Q}, \mathbb{Q}^+, \mathbb{R}, \mathbb{R}^+\}$. A rational stochastic language is a probability distribution over a free monoid Σ^* which is rational over K, that is which can be generated by a multiplicity automata with parameters in K. We study the relations between the classes of rational stochastic languages $\mathcal{S}_K^{rat}(\Sigma)$. We define the notion of *residual* of a stochastic language and we use it to investigate properties of several subclasses of rational stochastic languages. Lastly, we study the representation of rational stochastic languages by means of multiplicity automata.

1 Introduction

In probabilistic grammatical inference, data often arise in the form of a finite sequence of words $w_1, \ldots, w_n$ over some predefined alphabet Σ. These words are assumed to be independently drawn according to a fixed but unknown probability distribution over Σ^*. Probability distributions over free monoids Σ^* are called *stochastic languages*. A usual goal in grammatical inference is to try to infer an approximation of this distribution in some class of probabilistic models, such as *probabilistic automata*. A probabilistic automaton (PA) is composed of a *structure*, which is a finite automaton (NFA), and *parameters* associated with states and transitions, which represent the probability for a state to be initial, terminal or the probability for a transition to be chosen. It can easily be shown that probabilistic automata have the same expressivity as Hidden Markov Models (HMM), which are heavily used in statistical inference [DDE05]. Given the structure A of a probabilistic automaton and a sequence of words S, computing parameters for A which maximize the likelihood of S is NP-hard [AW92]. In practical cases however, algorithms based on the E.M. (*Expectation-Maximization*) method [DLR77] can be used to compute approximate values. On the other hand, inferring a probabilistic automaton (structure and parameters) from a sequence of words is a widely open field of research. Most results obtained so far only deal with restricted subclasses of PA, such as Probabilistic Deterministic Automata (PDA), i.e. probabilistic automata whose structure is deterministic (DFA) or Probabilistic Residual Automata (PRA), i.e. probabilistic automata whose structure is a residual finite state automaton (RFSA)[CO94,CO99,dlHT00,ELDD02,DE04].

In other respects, it can be noticed that stochastic languages are particular cases of *formal power series* and that probabilistic automata are also particular cases of *multiplicity automata*, notions which have been extensively studied in the field of formal language theory[SS78,BR84,Sak03]. Therefore, stochastic languages which can be generated by multiplicity automata are special cases of *rational languages*. We call them *rational stochastic languages*. The goal of the present paper is to provide a systematic and comprehensive study of *rational stochastic languages* so as to bring out

properties that could be useful for a grammatical inference purpose. Indeed, considering the objects to infer as special cases of rational languages makes it possible to use the powerful theoretical tools that have been developed in that field and hence, give answers to many questions that naturally arise when working with them: is it possible to decide within polynomial time whether two probabilistic automata generate the same stochastic language? does allowing negative coefficients in probabilistic automata extend the class of generated stochastic languages? can a rational stochastic language which takes all its values in $\mathbb{Q}$ always be generated by a multiplicity automata with coefficients in $\mathbb{Q}$? and so forth. Also, studying *rational stochastic languages* for themselves, considered as objects of language theory, helps to bring out notions and properties which are important in a grammatical inference pespective: for example, we show that the notion of residual language (or derivative), so important for grammatical inference [DLT02,DLT04], has a natural counterpart for stochastic languages [DE03], which can be used to express many properties of classes of stochastic languages.

Formal power series take their values in a semiring K: let us denote by $K\langle\langle\Sigma\rangle\rangle$ the set of all formal power series. Here, we only consider semirings $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{Q}^+$ and $\mathbb{R}^+$. For any such semiring K, we define the set $\mathcal{S}_K^{rat}(\Sigma)$ of rational stochastic languages as the set of stochastic languages over Σ which are rational languages over K. For any two distinct semirings K and K', the corresponding sets of rational stochastic languages are distinct. We show that $\mathbb{R}$ is a Fatou extension of $\mathbb{Q}$ for stochastic languages, which means that any rational stochastic language over $\mathbb{R}$ which takes its values in $\mathbb{Q}$ is also rational over $\mathbb{Q}$. However, $\mathbb{R}^+$ is not a Fatou extension of $\mathbb{Q}^+$ for stochastic languages: there exists a rational stochastic language over $\mathbb{R}^+$ which takes its values in $\mathbb{Q}^+$ and which is not rational over $\mathbb{Q}^+$.

For any stochastic language p over Σ and any word u such that $p(u\Sigma^*) \neq 0$, let us define the residual language $u^{-1}p$ of p with respect to u by $u^{-1}p(w) = p(uw)/p(u\Sigma^*)$: residual languages clearly are stochastic languages. We show that the residual languages of a rational stochastic language p over K are also rational over K. The residual subsemimodule $[Res(p)]$ of $K\langle\langle\Sigma\rangle\rangle$ spanned by the residual languages of any stochastic language p may be used to express the rationality of p: p is rational iff $[Res(p)]$ is included in a finitely generated subsemimodule of $K\langle\langle\Sigma\rangle\rangle$. But when K is positive, i.e. $K = \mathbb{Q}^+$ or $K = \mathbb{R}^+$, it may happen that $[Res(p)]$ itself is not finitely generated. We study the properties of two subclasses of $\mathcal{S}_K^{rat}(\Sigma)$: the set $\mathcal{S}_K^{fingen}(\Sigma)$ composed of rational stochastic languages over K whose residual subsemimodule is finitely generated and the set $\mathcal{S}_K^{fin}(\Sigma)$ composed of rational stochastic languages over K which have finitely many residual languages. We show that for any of these two classes, $\mathbb{R}^+$ is a Fatou extension of $\mathbb{Q}^+$: any stochastic language of $\mathcal{S}_{\mathbb{R}^+}^{fingen}(\Sigma)$ (resp. of $\mathcal{S}_{\mathbb{R}^+}^{fin}(\Sigma)$) which takes its values in $\mathbb{Q}^+$ is an element of $\mathcal{S}_{\mathbb{Q}^+}^{fingen}(\Sigma)$ (resp. of $\mathcal{S}_{\mathbb{Q}^+}^{fin}(\Sigma)$). We also show that for any element p of $\mathcal{S}_K^{fingen}(\Sigma)$, there exists a unique minimal subset of residual languages of p which generates $[Res(p)]$.

Then, we study the representation of rational stochastic languages by means of multiplicity automata. We first show that the set of multiplicity automata with parameters in $\mathbb{Q}$ which generate stochastic languages is not recursive. Moreover, it contains no recursively enumerable subset capable to generate the whole set of rational stochastic languages over $\mathbb{Q}$. A stochastic language p is a formal series which has two properties: (i) $p(w) \geq 0$ for any word w, (ii) $\sum_w p(w) = 1$. We show that the undecidability

comes from the first requirement, since the second one can be decided within polynomial time. We show that the set of stochastic languages which can be generated by probabilistic automata with parameters in $\mathbb{Q}^+$ (resp.$\mathbb{R}^+$) exactly coincides with $\mathcal{S}^{rat}_{\mathbb{Q}^+}(\Sigma)$ (resp. $\mathcal{S}^{rat}_{\mathbb{R}^+}(\Sigma)$). A probabilistic automaton A is called a Probabilistic Residual Automaton (PRA) if the stochastic languages associated with its states are residual languages of the stochastic languages p_A generated by A. We show that the set of stochastic languages that can be generated by probabilistic residual automata with parameters in $\mathbb{Q}^+$ (resp.$\mathbb{R}^+$) exactly coincides with $\mathcal{S}^{fingen}_{\mathbb{Q}^+}(\Sigma)$ (resp. $\mathcal{S}^{fingen}_{\mathbb{R}^+}(\Sigma)$). We do not know whether the class of PRA is decidable. However, we describe two decidable subclasses of PRA capable of generating $\mathcal{S}^{fingen}_{K}(\Sigma)$ when $K = \mathbb{Q}^+$ or $K = \mathbb{R}^+$: the class of K-reduced PRA and the class of prefixial PRA. The first one provides minimal representation in the class of PRA but we show that the membership problem is PSPACE-complete. The second one produces more cumbersome representation but the membership problem is polynomial. Finally, we show that the set of stochastic languages that can be generated by probabilistic deterministic automata with parameters in $\mathbb{Q}^+$ (resp.$\mathbb{R}^+$) exactly coincides with $\mathcal{S}^{fin}_{\mathbb{Q}^+}(\Sigma)$, which is also equal to $\mathcal{S}^{fin}_{\mathbb{Q}}(\Sigma)$ (resp. $\mathcal{S}^{fin}_{\mathbb{R}^+}(\Sigma)$, which is also equal to $\mathcal{S}^{fin}_{\mathbb{R}}(\Sigma)$).

We recall some properties on rational series, stochastic languages and multiplicity automata in Section 2. We define and study rational stochastic languages in Section 3. The relations between the classes of rational stochastic languages are studied in Subsection 3.1. Properties of the residual languages of rational stochastic languages are studied in Subsection 3.2. A characterisation of rational stochastic languages in terms of stable subsemimodule is given in Subsection 3.3. Classes $\mathcal{S}^{fingen}_{K}(\Sigma)$ and $\mathcal{S}^{fin}_{K}(\Sigma)$ are defined and studied in Subsection 3.4. The representation of rational stochastic languages by means of multiplicity automata is given in Section 4.

2 Preliminaries

2.1 Rational series

In this section, we recall some definitions and results on rational series. For more information, we invite the reader to consult [SS78,BR84,Sak03].

Let Σ be a finite *alphabet*, and Σ^* be the set of words on Σ. The empty word is denoted by ε and the length of a word u is denoted by $|u|$. The number of occurrences of the letter x in the word w is denoted by $|w|_x$. For any integer k, we denote by Σ^k the set $\{u \in \Sigma^* \mid |u| = k\}$ and by $\Sigma^{\leq k}$ the set $\{u \in \Sigma^* \mid |u| \leq k\}$. We denote by $<$ the length-lexicographic order on Σ^*. For any word $u \in \Sigma^*$ and any language $L \subseteq \Sigma^*$, let $uL = \{uv \in \Sigma^* | v \in L\}$ and $u^{-1}L = \{v \in \Sigma^* | uv \in L\}$. A subset P of Σ^* is *prefixial* if for any $u, v \in \Sigma^*$, $uv \in P \Rightarrow u \in P$.

A *semiring* is a set K with two binary operations $+$ and $\cdot$ and two constant elements 0 and 1 such that

1. $\langle K, +, 0\rangle$ is a commutative monoid,
2. $\langle K, \cdot, 1\rangle$ is a monoid,
3. the distribution laws $a \cdot (b + c) = a \cdot b + a \cdot c$ and $(a + b) \cdot c = a \cdot c + b \cdot c$ hold,
4. $0 \cdot a = a \cdot 0 = 0$ for every a.

A semiring is *positive* if the sum of two elements different from 0 is different from 0.

The semirings we consider here are the field of rational numbers $\mathbb{Q}$, the field of real numbers $\mathbb{R}$, $\mathbb{Q}^+$ and $\mathbb{R}^+$, respectively the non negative elements of $\mathbb{Q}$ and $\mathbb{R}$; $\mathbb{Q}^+$ and $\mathbb{R}^+$ are positive semirings.

Let Σ be a finite alphabet and K a semiring. A *formal power series* is a mapping r of Σ^* into K. The values $r(w)$ where $w \in \Sigma^*$ are referred to as the *coefficients* of the series, and r is written as a formal sum $r = \sum_{w\in\Sigma^*} r(w)w$. The set of all formal power series is denoted by $K\langle\langle\Sigma\rangle\rangle$. Given r, the subset of Σ^* defined by $\{w|r(w) \neq 0\}$ is the *support* of r and denoted by $supp(r)$. A *polynomial* is a series whose support is finite. The subset of $K\langle\langle\Sigma\rangle\rangle$ consisting of all polynomials is denoted by $K\langle\Sigma\rangle$.

We denote by 0 the series all of whose coefficients equal 0. We denote by 1 the series whose coefficient for ε equals 1, the remaining coefficients being equal to 0. The *sum* of two series r and r' in $K\langle\langle\Sigma\rangle\rangle$ is defined by $r + r' = \sum_{w\in\Sigma^*}(r(w) + r'(w))w$. The multiplication of a series r by a scalar $a \in K$ is defined by $ar = \sum_{w\in\Sigma^*} a \cdot r(w)w$. The Cauchy product of two series r and r' is defined by $rr' = \sum_{w\in\Sigma^*} \left(\sum_{w_1w_2=w} r(w_1) \cdot r'(w_2)\right) w$. These operations furnish $K\langle\langle\Sigma\rangle\rangle$ with the structure of a semiring with $K\langle\Sigma\rangle$ as a subsemiring. The Hadamard product of two series r and r' is defined by $r \odot r' = \sum_{w\in\Sigma^*} r(w)r'(w)w$.

A series r is *quasiregular* if $r(\epsilon) = 0$. Quasiregular series have the property that for every $w \in \Sigma^*$, there exist finitely many integers i such that $r^i(w) \neq 0$ where the exponent i of r^i refers to the Cauchy product. Let r be a quasiregular series, r^* (resp. r^+) is defined by $r^*(w) = \sum_{i\geq 0} r^i(w)$ (resp. $r^+(w) = \sum_{i\geq 1} r^i(w)$).

A subsemiring R of $K\langle\langle\Sigma\rangle\rangle$ is *rationally closed* if $r^+ \in R$ for every quasiregular element r of R. The family $K^{rat}\langle\langle\Sigma\rangle\rangle$ of K-rational series over Σ is the smallest rationally closed subset of $K\langle\langle\Sigma\rangle\rangle$ which contains all polynomials. When K is commutative, the Hadamard product of two rational series is a rational series.

Let K be a semiring and let m, n be two integers. Let us denote by $K^{m\times n}$ the set of $m \times n$ matrices whose elements belong to K and by I_m the matrix whose diagonal elements are equal to 1 and whose all other elements are null. Note that $K^{m\times m}$ forms a semiring.

A series r is *recognizable* if there exists a multiplicative homomorphism $\mu : \Sigma^* \to K^{n\times n}, n \geq 1$, and two matrices $\lambda \in K^{1\times n}, \gamma \in K^{n\times 1}$ such that for every $w \in \Sigma^*$, $r(w) = \lambda\mu(w)\gamma$. The tuple (λ, μ, γ) is called an n dimensional *linear representation* of r. A linear representation of r is said to be *reduced* if its dimension is minimal.

Let us denote by $K^{rec}\langle\langle\Sigma\rangle\rangle$ the set of all recognizable series.

Theorem 1. *[Sch61] The families $K^{rat}\langle\langle\Sigma\rangle\rangle$ and $K^{rec}\langle\langle\Sigma\rangle\rangle$ coincide.*

Let K be a semiring. Then a commutative monoid V is called a *K-semimodule* if there is an operation $\cdot$ from $K \times V$ into V such that for any $a, b \in K, v, w \in V$,

1. $(ab) \cdot v = a \cdot (b \cdot v)$,
2. $(a + b) \cdot v = a \cdot v + b \cdot v$ and $a \cdot (v + w) = a \cdot v + a \cdot w$,
3. $1 \cdot v = v$ and $0 \cdot v = 0$.

If S is a subset of a K-semimodule V, the subsemimodule $[S]$ generated by S is the smallest of all subsemimodules of V containing S. It can be proved that $[S] = \{a_1s_1 + \ldots + a_ns_n|n \in \mathbb{N}^*, a_i \in K, s_i \in S\}$.

Let us consider the semimodule K^{Σ^*} of all functions $F : \Sigma^* \to K$. For any word u of Σ^* and any function F of K^{Σ^*}, we define a new function $\dot{u}F$ by $\dot{u}F(v) = F(uv)$ for any word v. The operator transforming F into $\dot{u}F$ is linear: for any $F, G \in K^{\Sigma^*}$ and $a \in K$, $\dot{u}\,(a \cdot F) = a \cdot \dot{u}F$ and $\dot{u}(F + G) = \dot{u}F + \dot{u}G$. A subset B of K^{Σ^*} is called *stable* if the conditions $u \in \Sigma^*$ and $F \in B$ imply that $\dot{u}F \in B$.

Theorem 2. *[Fli74,Jac75] Suppose that K is a commutative semiring and r belongs to $K\langle\langle\Sigma\rangle\rangle$. Then the following three conditions are equivalent:*

1. *r belongs to $K^{rat}\langle\langle\Sigma\rangle\rangle$;*
2. *the subsemimodule of $K\langle\langle\Sigma\rangle\rangle$ generated by $\{\dot{u}r | u \in \Sigma^*\}$ is contained in a finitely generated stable subsemimodule of K^{Σ^*};*
3. *r belongs to a finitely generated stable subsemimodule of K^{Σ^*}.*

When K is not a field, it may happen that a series r belongs to a finitely generated stable subsemimodule of $K\langle\langle\Sigma\rangle\rangle$, and hence is a rational series, while the stable subsemimodule generated by $\{\dot{u}r | u \in \Sigma^*\}$ is not finitely generated. An example of this situation will be provided on Example 1.

Two linear representations (λ, μ, γ) and $(\lambda', \mu', \gamma')$ of a rational series r are *similar* if there exists an inversible matrix $m \in K^{n\times n}$ such that $\lambda' = \lambda m, \mu' w = m^{-1}\mu w m$ for any word w and $\gamma' = m^{-1}\gamma$.

Theorem 3. *[Sch61,Fli74] Assume that K is a commutative field. Then any two reduced linear representations (λ, μ, γ) and $(\lambda', \mu', \gamma')$ of a rational series r are similar. The dimension of any reduced linear representation of r is also the dimension of the vector subspace generated by $\{\dot{u}r | u \in \Sigma^*\}$.*

Let K be a subsemiring of K'. K' is said to be a *Fatou extension* of K if every rational series over K' with coefficients in K is a rational series over K. It has been shown in [Fli74] that when K and K' are commutative fields then K' is a Fatou extension of K. Therefore, $\mathbb{R}$ is a Fatou extension of $\mathbb{Q}$: any rational series over $\mathbb{R}$ which only takes rational values is a rational series over $\mathbb{Q}$: $\mathbb{R}^{rat}\langle\langle\Sigma\rangle\rangle \cap \mathbb{Q}\langle\langle\Sigma\rangle\rangle = \mathbb{Q}^{rat}\langle\langle\Sigma\rangle\rangle$. It has also been proved that $\mathbb{R}^+$ is not a Fatou extension of $\mathbb{Q}^+$: $\mathbb{Q}^{+rat}\langle\langle\Sigma\rangle\rangle \subsetneq \mathbb{R}^{+rat}\langle\langle\Sigma\rangle\rangle \cap \mathbb{Q}^+\langle\langle\Sigma\rangle\rangle$.

2.2 Stochastic languages

A *stochastic language* is a formal series p which takes its values in $\mathbb{R}^+$ and such that $\sum_{w\in\Sigma^*} p(w) = 1$. For any stochastic language p and any language $L \subseteq \Sigma^*$, the sum $\sum_{w\in L} p(w)$ is defined without ambiguity. So, let us denote $\sum_{w\in L} p(w)$ by $p(L)$. The set of all stochastic languages over Σ is denoted by $\mathcal{S}(\Sigma)$. For any stochastic language p and any word u such that $p(u\Sigma^*) \neq 0$, we define the stochastic language $u^{-1}p$ by

$$u^{-1}p(w) = \frac{p(uw)}{p(u\Sigma^*)}.$$

$u^{-1}p$ is called the *residual language* of p wrt u. Let us denote by $res(p)$ the set $\{u \in \Sigma^* | \sum_{w\in\Sigma^*} p(uw) \neq 0\}$ and by $Res(p)$ the set $\{u^{-1}p | u \in res(p)\}$. For any $K \in \{\mathbb{R}, \mathbb{R}^+, \mathbb{Q}, \mathbb{Q}^+\}$, define $\mathcal{S}_K^{rat}(\Sigma) = K^{rat}\langle\langle\Sigma\rangle\rangle \cap \mathcal{S}(\Sigma)$, the set of rational stochastic

languages over K. Let $S = \{s_1, \ldots, s_n\}$ be a finite subset of $\mathcal{S}(\Sigma)$. The convex hull of S in $K\langle\langle\Sigma\rangle\rangle$ is defined by $conv_K(S) = \{s \in K\langle\langle\Sigma\rangle\rangle | s = \alpha_1 \cdot s_1 + \ldots + \alpha_n \cdot s_n$ where each $\alpha_i \in K, \alpha_i \geq 0$ and $\alpha_1 + \ldots + \alpha_n = 1\}$. Clearly, any element of $conv_K(S)$ is a stochastic language.

Example 1. Let $\Sigma = \{a\}$, and let p_1, p_2 and p be the rational stochastic languages over $\mathbb{R}^+$ defined on Σ^* by

$$p_1(a^n) = 2^{-(n+1)}, p_2(a^n) = 3 \cdot 2^{-(2n+2)} \text{ and } p = (p_1 + p_2)/2.$$

Check that

$$\overline{\dot{a^n}} p_1 = \frac{p_1}{2^n}, \overline{\dot{a^n}} p_2 = \frac{p_2}{2^{2n}} \text{ and } \overline{\dot{a^n}} p = \frac{2^n p_1 + p_2}{2^{2n+1}}$$

and

$$(a^n)^{-1} p_1 = p_1, (a^n)^{-1} p_2 = p_2 \text{ and } (a^n)^{-1} p = \frac{2^n p_1 + p_2}{2^n + 1}.$$

Let $\mathcal{V}$ be the vector subspace of $\mathbb{R}\langle\langle\Sigma\rangle\rangle$ generated by p_1 and p_2: $\mathcal{V}$ is represented on Figure 1. The subsemimodule of $\mathbb{R}^+\langle\langle\Sigma\rangle\rangle$ generated by p_1 and p_2 corresponds

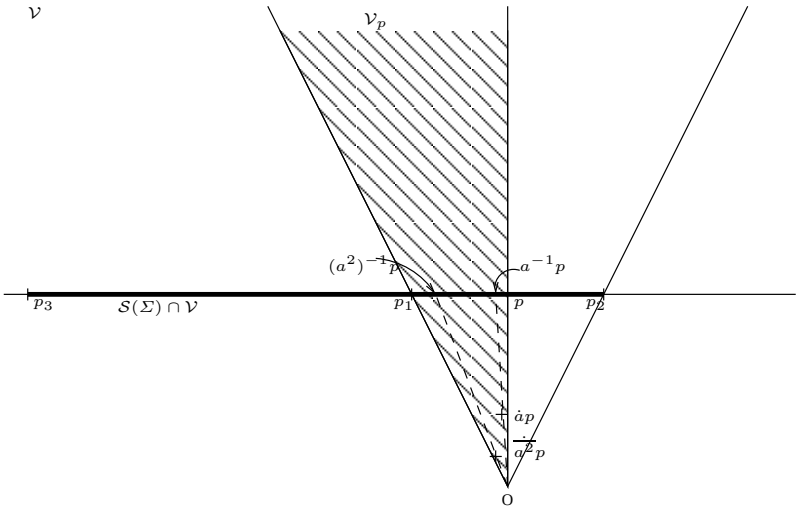

Fig. 1. The stable subsemimodule of $\mathbb{R}^+\langle\langle\Sigma\rangle\rangle$ generated by p is equal to $\mathcal{V}_p$: it does not contains the halfline $]Op_1)$ and it is not finitely generated.

to the closed halfcone $\mathcal{C}$ delimited by the halflines $[Op_1)$ and $[Op_2)$. The line (p_1p_2) is composed of the rational series r in $\mathcal{V}$ which satisfy $\sum_{w\in\Sigma^*} r(w) = 1$. Let $q = \alpha p_1 + (1-\alpha)p_2$. The constraint $q(a^n) \geq 0$ is equivalent to the inequality

$$(2^{n+1} - 3)\alpha + 3 \geq 0.$$

The series q such that $q(a^n) \geq 0$ for any integer n must satisfy

$$0 \leq \alpha \leq 3.$$

Let $p_3 = 3p_1 - 2p_2$. The stochastic languages in $\mathcal{V}$ are the points of the line (p_2p_3) which lie between p_2 and p_3.

Let $\mathcal{V}_p$ be the subsemimodule of $\mathbb{R}^+\langle\langle\Sigma\rangle\rangle$ generated by $\{\dot{u}p | u \in \Sigma^*\}$. Check that $\mathcal{V}_p = \{t(\alpha p_1 + (1-\alpha)p_2) | 1/2 \leq \alpha < 1, t \in \mathbb{R}^+\}$ and that $\mathcal{V}_p$ is not finitely generated.

2.3 Automata

A *non deterministic finite automaton* (NFA) is a tuple $\langle \Sigma, Q, Q_I, Q_T, \delta \rangle$ where Q is a finite set of states, $Q_I \subseteq Q$ is the set of initial states, $Q_T \subseteq Q$ is the set of final states, δ is the *transition function* defined from $Q \times \Sigma$ to 2^Q. Let δ also denote the extended transition function defined from $2^Q \times \Sigma^*$ to 2^Q by $\delta(q, \varepsilon) = \{q\}$, $\delta(q, wx) = \cup_{q' \in \delta(q,w)} \delta(q', x)$ and $\delta(R, w) = \cup_{q \in R} \delta(q, w)$ for any $q \in Q$, $R \subseteq Q$, $x \in \Sigma$ and $w \in \Sigma^*$. An NFA is *deterministic* (DFA) if Q_I contains only one element q_0 and if $\forall q \in Q, \forall x \in \Sigma, |\delta(q, x)| \leq 1$.

Let K be a semiring. A *K-multiplicity automaton (MA)* is a 5-tuple $\langle \Sigma, Q, \varphi, \iota, \tau \rangle$ where Q is a finite set of states, $\varphi : Q \times \Sigma \times Q \to K$ is the transition function, $\iota : Q \to K$ is the initialization function and $\tau : Q \to K$ is the termination function. Let $Q_I = \{q \in Q | \iota(q) \neq 0\}$ be the set of *initial states* and $Q_T = \{q \in Q | \tau(q) \neq 0\}$ be the set of *terminal states*. The *support* of an MA $\langle \Sigma, Q, \varphi, \iota, \tau \rangle$ is the NFA $\langle \Sigma, Q, Q_I, Q_T, \delta \rangle$ where $\delta(q, x) = \{q' \in Q | \varphi(q, x, q') \neq 0\}$. We extend the transition function φ to $Q \times \Sigma^* \times Q$ by $\varphi(q, wx, r) = \sum_{s \in Q} \varphi(q, w, s)\varphi(s, x, r)$ and $\varphi(q, \varepsilon, r) = 1$ if $q = r$ and 0 otherwise, for any $q, r \in Q$, $x \in \Sigma$ and $w \in \Sigma^*$. For any finite subset $L \subset \Sigma^*$ and any $R \subseteq Q$, define $\varphi(q, L, R) = \sum_{w \in L, r \in R} \varphi(q, w, r)$.

For any MA $A = \langle \Sigma, Q, \varphi, \iota, \tau \rangle$, we define the series r_A by

$$r_A(w) = \sum_{q,r \in Q} \iota(q)\varphi(q, w, r)\tau(r).$$

For any $q \in Q$, we define the series $r_{A,q}$ by $r_{A,q}(w) = \sum_{r \in Q} \varphi(q, w, r)\tau(r)$.

If the semiring K is positive, it can be shown that the support of the series r_A defined by a K-multiplicity automaton is equal to the language defined by the support of A. In particular, $supp(r_A)$ is a regular language. This property is false in general when K is not positive.

Two MA A and A' are *equivalent* if they define the same series, i.e. if $r_A = r_{A'}$.

Let $A = \langle \Sigma, Q, \varphi, \iota, \tau \rangle$ be a K-MA and let $q \in Q$. Suppose that there exist coefficients $\alpha_{q'} \in K$ for $q' \in Q' = Q \setminus \{q\}$ such that $r_{A,q} = \sum_{q' \in Q'} \alpha_{q'} r_{A,q'}$. Let $A' = \langle \Sigma, Q', \varphi', \iota', \tau' \rangle$ where

- $\varphi'(r, x, s) = \varphi(r, x, s) + \alpha_s \varphi(r, x, q)$ for any $r, s \in Q'$ and $x \in \Sigma$,
- $\iota'(r) = \iota(r) + \alpha_r \iota(q)$ for any $r \in Q'$,
- $\tau'(r) = \tau(r)$ for any $r \in Q'$.

The multiplicity automaton A' is called a *K-reduction* of A. A multiplicity automaton A is called *K-reduced* if it has no K-reduction.

Proposition 1. *Let $A = \langle \Sigma, Q, \varphi, \iota, \tau \rangle$ be a K-MA and let $A' = \langle \Sigma, Q', \varphi', \iota', \tau' \rangle$ be a K-reduction of A. Then, for any state $q' \in Q'$, $r_{A',q'} = r_{A,q'}$. As a consequence, $r_{A'} = r_A$.*

Proof. Let $Q' = Q \setminus \{q\}$ and let $\alpha_{q'} \in K$ for any $q' \in Q'$ such that $r_{A,q} = \sum_{q' \in Q'} \alpha_{q'} r_{A,q'}$. For any state $r \in Q'$, we have

$$r_{A',r}(\varepsilon) = \tau'(r) = \tau(r) = r_{A,r}(\varepsilon).$$

Now, assume that for any word w of length $\leq k$ and any state $r \in Q'$ we have $r_{A',r}(w) = r_{A,r}(w)$. Let x be a letter, we have:

$$\begin{aligned} r_{A',r}(xw) &= \sum_{s \in Q'} \varphi'(r,x,s) r_{A',s}(w) = \sum_{s \in Q'} \left(\varphi(r,x,s) + \alpha_s \varphi(r,x,q)\right) r_{A,s}(w) \\ &= \sum_{s \in Q'} \varphi(r,x,s) r_{A,s}(w) + \varphi(r,x,q) \sum_{s \in Q'} \alpha_s r_{A,s}(w) \\ &= \sum_{s \in Q'} \varphi(r,x,s) r_{A,s}(w) + \varphi(r,x,q) r_{A,q}(w) \\ &= \sum_{s \in Q} \varphi(r,x,s) r_{A,s}(w) = r_{A,r}(xw). \end{aligned}$$

Hence, $r_{A',r} = r_{A,r}$ for any r of Q'. Moreover,

$$\begin{aligned} r_{A'} &= \sum_{s \in Q'} \iota'(s) r_{A,s} = \sum_{s \in Q'} \left(\iota(s) + \alpha_s \iota(q)\right) r_{A,s} \\ &= \sum_{s \in Q'} \iota(s) r_{A,s} + \iota(q) \sum_{s \in Q'} \alpha_s r_{A,s} = \sum_{s \in Q} \iota(s) r_{A,s} = r_A. \end{aligned}$$

□

A state $q \in Q$ is *accessible* (resp. *co-accessible*) if there exists $q_0 \in Q_I$ (resp. $q_t \in Q_T$) and $u \in \Sigma^*$ such that $\varphi(q_0, u, q) \neq 0$ (resp. $\varphi(q, u, q_t) \neq 0$). An MA is *trimmed* if all its states are accessible and co-accessible. Given an MA A, a trimmed MA equivalent to A can efficiently be computed from A.

From now, we only consider trimmed MA.

We shall consider several subclasses of multiplicity automata, defined as follows:

A *semi Probabilistic Automaton (semi-PA)* is an MA $\langle \Sigma, Q, \varphi, \iota, \tau \rangle$ such that ι, φ and τ take their values in $[0,1]$, such that $\sum_{q \in Q} \iota(q) \leq 1$ and for any state q, $\tau(q) + \varphi(q, \Sigma, Q) \leq 1$. Semi-PA generate rational series over $\mathbb{R}^+$.

A *Probabilistic Automaton (PA)* is a trimmed semi-PA $\langle \Sigma, Q, \varphi, \iota, \tau \rangle$ such that $\sum_{q \in Q} \iota(q) = 1$ and for any state q, $\tau(q) + \varphi(q, \Sigma, Q) = 1$. Probabilistic automata generate stochastic languages.

Proposition 2. *Let $A = \langle \Sigma, Q, \varphi, \iota, \tau \rangle$ be a K-semi-PA (resp. a K-PA). For $q \in Q$, $\sum_{w \in \Sigma^*} r_{A,q}(w) \leq 1$ (resp. $\sum_{w \in \Sigma^*} r_{A,q}(w) = 1$). As a consequence, $\sum_{w \in \Sigma^*} r_A(w) \leq 1$ (resp. $\sum_{w \in \Sigma^*} r_A(w) = 1$).*

Proof. For any integer k and any $q \in Q$, we have

$$\begin{aligned}
&\sum_{|w|\leq k+1} r_{A,q}(w) + \varphi(q, \Sigma^{k+2}, Q) \\
&= \sum_{|w|\leq k} r_{A,q}(w) + \sum_{r\in Q} \varphi(q, \Sigma^{k+1}, r)\tau(r) + \sum_{r\in Q} \varphi(q, \Sigma^{k+1}, r)\varphi(r, \Sigma, Q) \\
&= \sum_{|w|\leq k} r_{A,q}(w) + \sum_{r\in Q} \varphi(q, \Sigma^{k+1}, r)[\tau(r) + \varphi(r, \Sigma, Q)].
\end{aligned}$$

From this relation, it is easy to infer by induction on k that

$$\sum_{|w|\leq k} r_{A,q}(w) + \sum_{r\in Q} \varphi(q, \Sigma^{k+1}, r) \leq 1 \text{ (resp. } = 1)$$

when A is a semi-PA (resp. a PA).

A first consequence is that

$$\sum_{w\in\Sigma^*} r_{A,q}(w) \leq 1 \text{ and } \sum_{w\in\Sigma^*} r_A(w) = \sum_{w\in\Sigma^*}\sum_{q\in Q} \iota(q) r_{A,q}(w) \leq 1.$$

Let $n = |Q|$. Since A is trimmed, there exists a word $u \in \Sigma^{\leq n-1}$ such that $r_{A,q}(u) > 0$. Therefore, there exists $\alpha < 1$ such that $\varphi(q, \Sigma^n, Q) < \alpha$. It can easily be shown, by induction on the integer k, that $\varphi(q, \Sigma^{kn}, Q) < \alpha^k$.

Now, when A is a PA, we have

$$\sum_{w\in\Sigma^*} r_{A,q}(w) \geq \sum_{|w|<kn} r_{A,q}(w) = 1 - \varphi(q, \Sigma^{kn}, Q) > 1 - \alpha^k$$

for any integer k. Therefore,

$$\sum_{w\in\Sigma^*} r_{A,q}(w) = 1.$$

Finally,

$$\sum_{w\in\Sigma^*} r_A(w) = \sum_{w\in\Sigma^*}\sum_{q\in Q} \iota(q) r_{A,q}(w) = \sum_{q\in Q} \iota(q) = 1.$$

□

It can easily be deduced from Proposition 2 that a $\mathbb{R}^+$-reduction of a PA is still a PA (the property is false in general for a semi-PA).

A *Probabilistic Residual Automaton (PRA)* is a PA $\langle \Sigma, Q, \varphi, \iota, \tau\rangle$ such that for any $q \in Q$, there exists a word u such that $r_{A,q} = u^{-1} r_A$. Check that a $\mathbb{R}^+$-reduction of a PRA is still a PRA, since the series associated with the states remain unchanged within a reduction.

A *Probabilistic Deterministic Automaton (PDA)* is a PA whose support is deterministic. Check that a PDA is a PRA. Therefore, a $\mathbb{R}^+$-reduction of a PDA is a PRA, but since reduction introduces non-determinism, it is no longer a PDA.

For any class C of K-multiplicity automata, let us denote by $\mathcal{S}_K^C(\Sigma)$ the class of all stochastic languages which are recognized by an element of C.

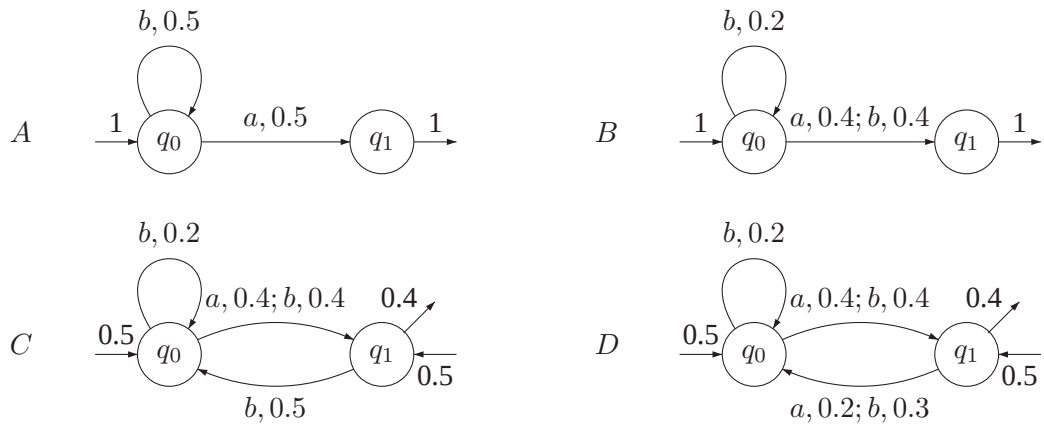

Fig. 2. Let us precise notations on automaton A: q_0 is the unique initial state and $\iota(q_0) = 1$, q_1 is the unique terminal state and $\tau(q_1) = 1$, $\varphi(q_0, a, q_1) = 0.5$, $\varphi(q_0, b, q_0) = 0.5$ and any other transitions satisfy $\varphi(q, x, q') = 0$. A is a PDA; B is a PRA since $r_{B,q_0} = r_B$ and $r_{B,q_1} = a^{-1}r_B$; C is also a PRA since $r_{C,q_0} = ab^{-1}r_C$ and $r_{C,q_1} = a^{-1}r_C$; it can easily be shown that D is not a PRA.

2.4 Equivalent representations of rational series

Stable finitely generated subsemimodules, linear representations and multiplicity automata provide us with several representations of rational series. The following classical claims show that they are equivalent: in particular, a series r over K is rational iff there exists a K-multiplicity automaton A such that $r = r_A$. Moreover, any one of these representations can efficiently be derived from any other one.

Claim 1 Let M be a stable subsemimodule of $K\langle\langle\Sigma\rangle\rangle$ generated by $r_1, \ldots, r_n$ and containing the series r. Let α_i and $\alpha_{i,j}^x$ be coefficients in K defined for any letter x and any $1 \leq i, j \leq n$ such that

$$r = \sum_{i=1}^{n} \alpha_i r_i \text{ and } \dot{x} r_i = \sum_{j=1}^{n} \alpha_{i,j}^x r_j.$$

Let (λ, μ, γ) be the linear representation defined by $\lambda[1, i] = \alpha_i$, $\mu(x)[i, j] = \alpha_{i,j}^x$ and $\gamma[i, 1] = r_i(\varepsilon)$ for any $1 \leq i, j \leq n$ and any $x \in \Sigma$. Then (λ, μ, γ) is a linear representation of r.

Claim 2 Let (λ, μ, γ) be an n-dimensional linear representation of r and let $A = \langle\Sigma, Q, \varphi, \iota, \tau\rangle$ be the MA defined by $Q = \{1, \ldots, n\}$, $\iota(i) = \lambda[1, i]$, $\tau(i) = \gamma[i, 1]$ and $\varphi(i, x, j) = \mu(x)[i, j]$. Then $r = r_A$.

Claim 3 Let $A = \langle\Sigma, Q, \varphi, \iota, \tau\rangle$ be an MA and let M be the subsemimodule generated by $\{r_{A,q} | q \in Q\}$. Then M is a stable subsemimodule of $K\langle\langle\Sigma\rangle\rangle$ which contains r_A.

The proofs of these claims are classical. We give them for sake of completness.

Proof (Claim 1). Let us prove by induction on the length of the word w that for any word w, $\mu(w)\gamma = (r_1(w), \ldots, r_n(w))^t$. From definition, $\mu(\varepsilon)\gamma = \gamma = (r_1(\varepsilon), \ldots, r_n(\varepsilon))^t$.

Suppose that the relation is proved for all words of length $\leq n$ and let $w \in \Sigma^n$ and $x \in \Sigma$.

$$\begin{aligned}
\mu(xw)\gamma &= \mu(x)\mu(w)\gamma \\
&= \mu(x)(r_1(w), \ldots, r_n(w))^t \text{ by induction hypothesis} \\
&= \left(\sum_{j=1}^{n} \alpha_{1,j}^{x} r_j(w), \ldots, \sum_{j=1}^{n} \alpha_{n,j}^{x} r_j(w) \right)^t \\
&= (\dot{x}r_1(w), \ldots, \dot{x}r_n(w))^t \\
&= (r_1(xw), \ldots, r_n(xw))^t .
\end{aligned}$$

Now, for any word w,

$$\lambda\mu(w)\gamma = \lambda(r_1(w), \ldots, r_n(w))^t = \sum_{i=1}^{n} \alpha_i r_i(w) = r(w).$$

□

Proof (Claim2). For any word w, we have

$$r_A(w) = \sum_{i,j=1}^{n} \iota(i)\varphi(i, w, j)\tau(j) = \sum_{i,j=1}^{n} = \lambda[1, i]\mu(w)[i, j]\gamma[i, 1] = \lambda\mu(w)\gamma.$$

□

Proof (Claim3). First note that $r_A = \sum_{q \in Q} \iota(q) r_{A,q}$ and therefore, $r_A \in M$.

Next, for any letter x, any word w and any state $q \in Q$,

$$\dot{x}r_{A,q}(w) = r_{A,q}(xw) = \sum_{q' \in Q} \varphi(q, x, q')r_{A,q'}(w)$$

and therefore,

$$\dot{x}r_{A,q} = \sum_{q' \in Q} \varphi(q, x, q')r_{A,q'}.$$

M is a stable subsemimodule of $K\langle\langle\Sigma\rangle\rangle$. □

These equivalent characterizations make it possible to transfer definitions from one representation mode to another: check that an n-dimensional linear representation of a rational series over K is reduced if and only iff the corresponding multiplicity automaton is K-reduced.Also, results obtained using one representation can immediatly be transfered to the other ones.

2.5 Computing equivalence and reduction of MA

Deciding whether two NFA are equivalent is a PSPACE-complete problem. However, deciding whether two MA are equivalent can be achieved within polynomial time.

Proposition 3. *It is decidable within polynomial time whether two MAs over $\mathbb{R}$ are equivalent.*

Proof. Let A and A' be two MA and let (λ, μ, γ) (resp. $(\lambda', \mu', \gamma')$) be an n-dimensional (resp. n'-dimensional) linear representation of the rational series r_A (resp. $r_{A'}$). For any word w let $\theta(w) = (\mu(w)\gamma, \mu'(w)\gamma')$. Let E be the vector subspace of $\mathbb{R}^{n+n'}$ spanned by $\{\theta(w)|w \in \Sigma^*\}$ and let T be the linear mapping from $\mathbb{R}^{n+n'}$ to $\mathbb{R}$ defined by $T(u, u') = \lambda u - \lambda' u'$ for any $u \in \mathbb{R}^n$ and $u' \in \mathbb{R}^{n'}$. The series r_A and $r_{A'}$ are equal, i.e. A and A' are equivalent, iff $\forall (u, u') \in E, T(u, u') = 0$, property which can be checked within polynomial time. □

The following algorithm decides the equivalence of two MA:

```
Input: A, A' MA
B = {ε}, S = {x|x ∈ Σ}
while S ≠ ∅ do
    let v be the smallest element in S and let S = S \ {v}
    if θ(v) does not belong to the subspace spanned by θ(B)
then
        B = B ∪ {v} and S = S ∪ {vx|x ∈ Σ}
    end if
end while
while B ≠ ∅ do
    let v ∈ B and let B = B \ {v}
    if T(θ(v)) ≠ 0 then
        output no ; exit
    end if
end while
output yes.
```

The first part of the algorithm computes a basis of E; the second part checks whether $T(E) = \{0\}$.

Note that when A and A' are not equivalent, the previous algorithm provides a word u such that $r_A(u) \neq r_{A'}(u)$ and whose length is $\leq |Q| + |Q'|$.

Proposition 4. *Let $A_0, A_1, \ldots, A_n$ be MAs over $\mathbb{R}$. It is decidable within polynomial time whether there exists $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ such that $r_{A_0} = \sum_{i=1}^n \alpha_i r_{A_i}$. More precisely, all such tuples of parameters $(\alpha_1, \ldots, \alpha_n)$ are solutions of a linear system computable within polynomial time.*

Proof. Consider the following algorithm.

```
Let Eq = {r_{A_0}(ε) = Σ_{i=1}^n x_i r_{A_i}(ε)}
#Eq is a set of independent equations on variables x_1, ..., x_n.
While Eq has a solution (α_1, ..., α_n) such that r_{A_0} ≠ Σ_{i=1}^n α_i r_{A_i}
```

Let u be a word such that $r_{A_0}(u) \neq \sum_{i=1}^{n} \alpha_i r_{A_i}(u)$
$Eq = Eq \cup \{r_{A_0}(u) = \sum_{i=1}^{n} x_i r_{A_i}(u)\}$
Output : Eq

From Proposition 3, if $r_{A_0} \neq \sum_{i=1}^{n} \alpha_i r_{A_i}$, a word u such that $r_{A_0}(u) \neq \sum_{i=1}^{n} \alpha_i r_{A_i}(u)$ and whose length is $\leq \sum_{i=0}^{n} |Q_i|$ can be found within polynomial time (where $|Q_i|$ is the number of states of A_i). The algorithms ends since Eq has at most $n+1$ elements. It is clear that $(\alpha_1, \ldots, \alpha_n)$ is a solution of Eq iff $r_{A_0} = \sum_{i=1}^{n} \alpha_i r_{A_i}$. □

A similar result holds when we ask for positive coefficients.

Proposition 5. *Let $A_0, A_1, \ldots, A_n$ be MAs over $\mathbb{R}$. It is decidable within polynomial time whether there exists $\alpha_1, \ldots, \alpha_n \in \mathbb{R}^+$ such that $r_{A_0} = \sum_{i=1}^{n} \alpha_i r_{A_i}$.*

Proof. Add the constraints $x_1 \geq 0, \ldots, x_n \geq 0$ to the system Eq in the previous algorithm. A polynomial linear programming algorithm will then find a solution of Eq or decide that Eq has no solution. □

As a consequence of these propositions, it can efficiently be decided whether an MA A is K-reduced .

Proposition 6. *Let $A = \langle \Sigma, Q, \varphi, \iota, \tau \rangle$ be a K-MA. It is decidable within polynomial time whether A is K-reduced; if A is not K-reduced, a K reduction can be computed within polynomial time.*

Proof. For any $q \in Q$, check whether there exist coefficients $\alpha_{q'} \in K$ for $q' \in Q' = Q \setminus \{Q\}$ such that $r_{A,q} = \sum_{q' \in Q'} \alpha_{q'} r_{A,q'}$. If so, use these coefficients to compute a K-reduction of A. □

3 Rational stochastic languages

The objects we study are *rational stochastic languages*, i.e. stochastic languages which are also rational series. A rational stochastic language can always be generated by using a multiplicity automaton. But depending on the set K of numbers used for the parameters, we obtain different sets $\mathcal{S}_K^{rat}(\Sigma)$ of rational stochastic languages. In the following, we suppose that $K \in \{\mathbb{R}, \mathbb{R}^+, \mathbb{Q}, \mathbb{Q}^+\}$. First, we study the relations between all these classes of rational stochastic languages and next, we give a characterization of $\mathcal{S}_K^{rat}(\Sigma)$ in terms of stable subsemimodules of $\mathcal{S}(\Sigma)$.

3.1 Relations between classes of rational stochastic languages

Let us begin by the simplest inclusions.

Proposition 7.

$$\mathcal{S}_{\mathbb{Q}^+}^{rat}(\Sigma) \subseteq \mathcal{S}_{\mathbb{Q}}^{rat}(\Sigma) \subsetneq \mathcal{S}_{\mathbb{R}}^{rat}(\Sigma) \text{ and } \mathcal{S}_{\mathbb{Q}^+}^{rat}(\Sigma) \subsetneq \mathcal{S}_{\mathbb{R}^+}^{rat}(\Sigma) \subseteq \mathcal{S}_{\mathbb{R}}^{rat}(\Sigma).$$

Moreover,

$$\mathcal{S}_{\mathbb{R}^+}^{rat}(\Sigma) \setminus \mathbb{Q}\langle\langle \Sigma \rangle\rangle \neq \emptyset.$$

Proof. Let K_1 be a subsemiring of K_2. We have $K_1^{rat}\langle\langle\Sigma\rangle\rangle \subseteq K_2^{rat}\langle\langle\Sigma\rangle\rangle$ and hence, $\mathcal{S}_{K_1}^{rat}(\Sigma) \subseteq \mathcal{S}_{K_2}^{rat}(\Sigma)$.

Now, let r be the rational series defined on $\Sigma = \{a\}$ by $r(\varepsilon) = \sqrt{2}/2, r(a) = 1 - \sqrt{2}/2$ and $r(a^n) = 0$ for any $n \geq 2$. Clearly, $r \in \mathcal{S}_{\mathbb{R}^+}^{rat}(\Sigma) \setminus \mathbb{Q}\langle\langle\Sigma\rangle\rangle$ which implies that $\mathcal{S}_{\mathbb{Q}}^{rat}(\Sigma) \subsetneq \mathcal{S}_{\mathbb{R}}^{rat}(\Sigma)$ and $\mathcal{S}_{\mathbb{Q}^+}^{rat}(\Sigma) \subsetneq \mathcal{S}_{\mathbb{R}^+}^{rat}(\Sigma)$. □

A rational stochastic language over $\mathbb{R}$ which only takes rational values is a rational stochastic language over $\mathbb{Q}$.

Proposition 8.

$$\mathcal{S}_{\mathbb{R}}^{rat}(\Sigma) \cap \mathbb{Q}\langle\langle\Sigma\rangle\rangle = \mathcal{S}_{\mathbb{Q}}^{rat}(\Sigma).$$

Proof.

Recall that $\mathbb{R}$ is a Fatou extension of $\mathbb{Q}$: any rational series over $\mathbb{R}$ which only takes rational values is a rational series over $\mathbb{Q}$ i.e.

$$\mathbb{R}^{rat}\langle\langle\Sigma\rangle\rangle \cap \mathbb{Q}\langle\langle\Sigma\rangle\rangle = \mathbb{Q}^{rat}\langle\langle\Sigma\rangle\rangle.$$

As a consequence,

$$\begin{aligned}\mathcal{S}_{\mathbb{R}}^{rat}(\Sigma) \cap \mathbb{Q}\langle\langle\Sigma\rangle\rangle &= \mathcal{S}(\Sigma) \cap \mathbb{R}^{rat}\langle\langle\Sigma\rangle\rangle \cap \mathbb{Q}\langle\langle\Sigma\rangle\rangle \\ &= \mathcal{S}(\Sigma) \cap \mathbb{Q}^{rat}\langle\langle\Sigma\rangle\rangle \\ &= \mathcal{S}_{\mathbb{Q}}^{rat}(\Sigma).\end{aligned}$$

□

It has also been proved that $\mathbb{R}^+$ is not a Fatou extension of $\mathbb{Q}^+$: ${\mathbb{Q}^+}^{rat}\langle\langle\Sigma\rangle\rangle \subsetneq {\mathbb{R}^+}^{rat}\langle\langle\Sigma\rangle\rangle \cap \mathbb{Q}^+\langle\langle\Sigma\rangle\rangle$. We prove below that this result can be extended to stochastic languages: there exists a rational stochastic language over $\mathbb{R}^+$ which takes only rational values and which is not a rational stochastique language over $\mathbb{Q}^+$.

Proposition 9. $\mathcal{S}_{\mathbb{Q}^+}^{rat}(\Sigma) \subsetneq \mathcal{S}_{\mathbb{R}^+}^{rat}(\Sigma) \cap \mathbb{Q}^+\langle\langle\Sigma\rangle\rangle$.

Proof. We use an element in ${\mathbb{R}^+}^{rat}\langle\langle\Sigma\rangle\rangle \cap \mathbb{Q}^+\langle\langle\Sigma\rangle\rangle \setminus {\mathbb{Q}^+}^{rat}\langle\langle\Sigma\rangle\rangle$ described in [BR84] to prove the proposition.

Consider the multiplicity automaton $A = \langle\Sigma, Q, \varphi, \iota, \tau\rangle$ where $\Sigma = \{a, b\}$, $Q = \{q_0, q_1\}$, $\iota(q_0) = \iota(q_1) = 1$, $\varphi(q_0, a, q_0) = \alpha^2$, $\varphi(q_0, b, q_0) = \alpha^{-2}$, $\varphi(q_1, a, q_1) = \alpha^{-2}$, $\varphi(q_1, b, q_1) = \alpha^2$ where $\alpha = (\sqrt{5}+1)/2$, $\varphi(q_i, x, q_j) = 0$ for any $x \in \Sigma$ when $i \neq j$ and $\tau(q_0) = \tau(q_1) = 1$ (see Figure 3).

Let r_A be the rational series generated by A. Let $w \in \Sigma^*$. We have $r_A(w) = \alpha^{2n} + \alpha^{-2n}$ where $n = |w|_a - |w|_b$. Check that for any integer n, $\alpha^{2n} + \alpha^{-2n} \in \mathbb{N}$. Hence, $r_A \in {\mathbb{R}^+}^{rat}\langle\langle\Sigma\rangle\rangle \cap \mathbb{Q}^+\langle\langle\Sigma\rangle\rangle$. It is shown in [BR84] that $r_A \notin {\mathbb{Q}^+}^{rat}\langle\langle\Sigma\rangle\rangle$.

Now let $A' = \langle\Sigma, Q, \varphi', \iota', \tau'\rangle$ where for any states q and q' and any letter x, $\iota'(q) = 1/2$, $\varphi'(q, x, q') = \varphi(q, x, q')/4$ and $\tau'(q_0) = \tau'(q_1) = 1/4$. Check that $\alpha^2 + \alpha^{-2} = 3$. Then, A' is a probabilistic automaton. Let p be the stochastic language generated by A. We have

$$p(w) = \frac{1}{2^{2|w|+3}}\left(\alpha^{2n} + \alpha^{-2n}\right) \text{ where } n = |w|_a - |w|_b$$

and hence

$$p \in \mathcal{S}^{rat}_{\mathbb{R}^+}(\Sigma) \cap \mathbb{Q}^+\langle\langle\Sigma\rangle\rangle.$$

Let s be the series defined by $s(w) = 2^{2|w|+3}$. Clearly, $s \in \mathbb{Q}^{+rat}\langle\langle\Sigma\rangle\rangle$ and $r_A = s \odot p$ (Hadamard product). Recall that when K is commutative, the Hadamard product of two rational series is a rational series. Therefore $r_A \notin \mathbb{Q}^{+rat}\langle\langle\Sigma\rangle\rangle \Rightarrow p \notin \mathbb{Q}^{+rat}\langle\langle\Sigma\rangle\rangle$ and hence, $p \notin \mathcal{S}^{rat}_{\mathbb{Q}^+}(\Sigma)$. □

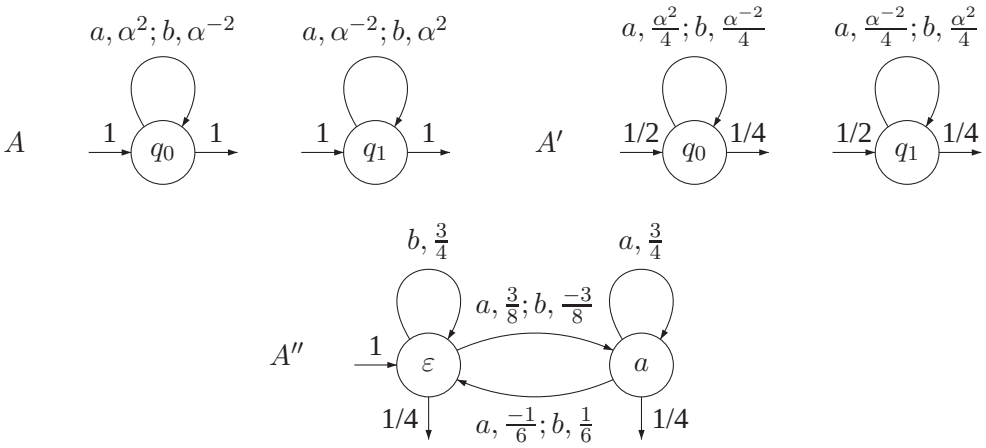

Fig. 3. A' generates a rational stochastic language $p_{A'}$ which takes all its values in $\mathbb{Q}$. However, $p_{A'}$ is not a rational stochastic language over $\mathbb{Q}^+$. A'' is a multiplicity automaton over $\mathbb{Q}$ which generates $p_{A'}$.

Remark that since p is a rational stochastic language which takes all its values in $\mathbb{Q}$, p is a rational stochastic language over $\mathbb{Q}$, from Prop 8. Let $p_0 = p_{A',q_0}$ and $p_1 = p_{A',q_1}$ be the stochastic languages generated from the states q_0 and q_1 of automaton A'. It can easily be shown that

$$\begin{cases} p = \frac{1}{2}p_0 + \frac{1}{2}p_1 \\ a^{-1}p = \frac{\alpha^2}{3}p_0 + \frac{\alpha^{-2}}{3}p_1 \end{cases}$$

These relations makes it possible to base on p and $a^{-1}p$ an automata which recognizes p. Check that

$$\dot{a}p = \frac{3}{8}p, \dot{b}p = \frac{3}{4}p - \frac{3}{8}a^{-1}p, \dot{a}a^{-1}p = \frac{-1}{6}p - \frac{3}{4}a^{-1}p \text{ and } \dot{b}a^{-1}p = \frac{1}{6}p + \frac{3}{4}a^{-1}p.$$

These relations can be used to prove that the automaton A'' in Fig. 3 generates p.

Now, we prove that there exists a rational stochastic language over $\mathbb{Q}$ which is not rational over $\mathbb{R}^+$. In particular, it cannot be generated by a probabilistic automaton.

Proposition 10. $\mathcal{S}^{rat}_{\mathbb{Q}}(\Sigma) \setminus \mathcal{S}^{rat}_{\mathbb{R}^+}(\Sigma) \neq \emptyset$.

Proof. Let $\Sigma = \{a, b\}$ and for any $w \in \Sigma^*$, let r and s be the series defined by $r(w) = |w|_a$ and $s(w) = |w|_b$. They are rational over $\mathbb{Q}$ since they belong to a stable finitely generated subsemimodule of $\mathbb{Q}\langle\langle\Sigma\rangle\rangle$. Indeed,

$$\dot{a}r = r + 1, \dot{b}r = r, \dot{a}s = s \text{ and } \dot{b}s = s + 1.$$

Hence, the series $r - s$ and $(r - s)^2$ where the exponent refers to the Hadamard product are also rational over $\mathbb{Q}$. For any $n \in \mathbb{N}$, let $\sigma_n = \sum_{w\in\Sigma^n}(r-s)^2(w) \leq n^2 \cdot 2^n$. Check that

$$\sigma_n = n2^n \text{ and } \sigma = \sum_{n\geq 0} \frac{\sigma_n}{2^{2n}} = 2.$$

Now, let t be the series defined by

$$t(w) = \frac{(r-s)^2(w)}{\sigma \cdot 2^{2|w|}}.$$

t is a rational stochastic languages over $\mathbb{Q}$. Its support is the set $supp(t) = \{w \in \Sigma^* \mid |w|_a \neq |w|_b\}$ which is known to be not rational. If t were rational over $\mathbb{R}^+$, it support would be rational. Therefore, $t \in \mathcal{S}^{rat}_{\mathbb{Q}}(\Sigma) \setminus \mathcal{S}^{rat}_{\mathbb{R}^+}(\Sigma)$. □

All these results can be summarized on diagram 4.

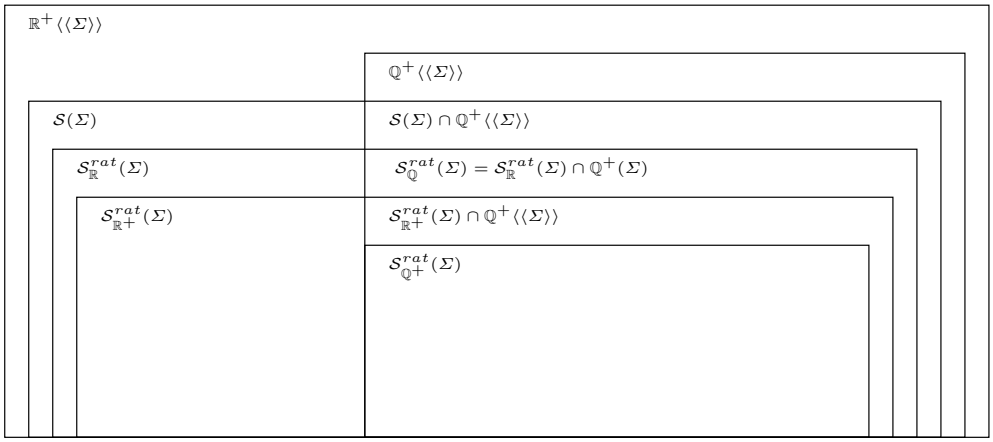

Fig. 4. Inclusion relations between classes of rational stochastic languages.

3.2 Residual languages of rational stochastic languages

Recall that given a stochastic language $p \in \mathcal{S}(\Sigma)$ and a word $u \in res(p)$, i.e. such that $p(u\Sigma^*) \neq 0$, the residual language of p wrt u is the stochastic language defined by

$$u^{-1}p(w) = \frac{p(uw)}{p(u\Sigma^*)}.$$

When p takes its values in $\mathbb{Q}^+$, it is not true in general that $u^{-1}p$ takes also its values in $\mathbb{Q}^+$.

Consider two series $(\alpha_n)_{n\in\mathbb{N}}$ and $(\beta_n)_{n\in\mathbb{N}}$ over $\mathbb{Q}^+$ and such that $\sum_n \alpha_n = \sqrt{2}/2$ and $\sum_n \beta_n = 4/5 - \sqrt{2}/2$. Now, consider the series $r \in \mathbb{Q}^+\langle\langle\{a,b\}\rangle\rangle$ defined by $r(\varepsilon) = 1/5$, $r(a^n) = \alpha_{n-1}$, $r(b^n) = \beta_{n-1}$ for $n \geq 1$ and $r(w) = 0$ otherwise. It is easy to check that r is a stochastic language which takes its values over $\mathbb{Q}^+$ and that $a^{-1}r(\varepsilon) = \sqrt{2}\alpha_0$. Therefore, $a^{-1}r \notin \mathbb{Q}\langle\langle\Sigma\rangle\rangle$.

We prove below that when p is a rational stochastic language over K, all its residual languages are also rational over K. Moreover, the set $Res(p) = \{u^{-1}p | u \in res(p)\}$ generates the same subsemimodule of $K\langle\langle\Sigma\rangle\rangle$ as the set $\{\dot{u}p | u \in \Sigma^*\}$.

We need before two linear algebra technical lemmas to prove this result.

Lemma 1. *Let $f : \mathbb{Q}^n \to \mathbb{Q}^n$ be a linear mapping and let $\mathbf{t} \in \mathbb{Q}^n$ such that $\sum_{k\geq 0} f^k\mathbf{t}$ converges to $\mathbf{u}$. Then $\mathbf{u} \in \mathbb{Q}^n$.*

Proof. Let F be the vector subspace of $\mathbb{Q}^n$ generated by $\{f^k\mathbf{t} | k \in \mathbb{N}\}$. There exists an integer d such that $f^0\mathbf{t} = \mathbf{t}, \ldots, f^{d-1}\mathbf{t}$ is a basis of F. As the sum $\sum_{k\geq 0} f^k\mathbf{t}$ converges, $f^k\mathbf{t}$ converges to 0 when k tends to infinity. Therefore, for any $\mathbf{v} \in F$, $f^k\mathbf{v}$ also converges to 0 when k tends to infinity. Let $\mathbf{v} \in F$ such that $f\mathbf{v} = \mathbf{v}$. We have also $f^k\mathbf{v} = \mathbf{v}$ for any integer k and hence, $\mathbf{v} = 0$. Let $g : F \to F$ defined by $g(\mathbf{v}) = \mathbf{v} - f\mathbf{v}$. The linear mapping g is one-to-one and for any $\mathbf{v} \in F$ and any integer k,

$$\mathbf{v} + f\mathbf{v} + \ldots + f^k\mathbf{v} = g^{-1}(1 - f^{k+1})(\mathbf{v}).$$

Therefore,

$$\mathbf{u} = g^{-1}\mathbf{t} \text{ and } \mathbf{u} \in \mathbb{Q}^n.$$

We use Lemma 1 to show that if $\{r_1, \ldots, r_n\}$ generates a stable subsemimodule of $\mathbb{Q}\langle\langle\Sigma\rangle\rangle$ and if each sum $\sum_{w\in\Sigma} r_i(w)$ converges to σ_i then each $\sigma_i \in \mathbb{Q}$.

Lemma 2. *Let M be a stable subsemimodule of $\mathbb{Q}\langle\langle\Sigma\rangle\rangle$ generated by $\{r_1, \ldots, r_n\}$ and let $\sigma_i^k = \sum_{w\in\Sigma^k} r_i(w)$ for any $1 \leq i \leq n$ and any integer k. Suppose that for any $1 \leq i \leq n$, the sums $\sum_{k\geq 0} \sigma_i^k$ converges to σ_i. Then $\sigma_i \in \mathbb{Q}$ for any $1 \leq i \leq n$.*

Proof. Let $\mathbf{t} = (r_1(\varepsilon), \ldots, (r_n(\varepsilon)))^t$. As M is stable, there exist $\alpha_{i,j}^x \in \mathbb{Q}$ for any $1 \leq i, j \leq n$ and any $x \in \Sigma$ such that $\dot{x}r_i = \sum_{j=1}^n \alpha_{i,j}^x \cdot r_j$. Let $B \in \mathbb{Q}^{n\times n}$ defined by $B[i,j] = \sum_{x\in\Sigma} \alpha_{i,j}^x$. Let us prove by induction on k that for any integer k, we have $(\sigma_1^k, \ldots, \sigma_n^k)^t = B^k\mathbf{t}$. The property is true for $k = 0$ as for any integer i, $\sigma_i^0 = r_i(\varepsilon)$.

Now,

$$\begin{aligned}
\sigma_i^{k+1} &= \sum_{w\in\Sigma^k, x\in\Sigma} r_i(xw) \\
&= \sum_{w\in\Sigma^k, x\in\Sigma} \dot{x}r_i(w) \\
&= \sum_{w\in\Sigma^k, x\in\Sigma, j\in\{1,\ldots,n\}} \alpha_{i,j}^x \cdot r_j(w) \\
&= \sum_{j\in\{1,\ldots,n\}} \left(\sum_{x\in\Sigma} \alpha_{i,j}^x\right) \cdot \sum_{w\in\Sigma^k} r_j(w) \\
&= \sum_{j\in\{1,\ldots,n\}} B[i,j]\sigma_j^k \\
&= \sum_{j\in\{1,\ldots,n\}} B[i,j](B^k\mathbf{t})[j] \text{ by induction hypothesis} \\
&= (B^{k+1}\mathbf{t})[i].
\end{aligned}$$

Therefore, $B^k\mathbf{t}$ converges to $(\sigma_1,\ldots,\sigma_n)^t$. From Lemma 1, $\sigma_i \in \mathbb{Q}$ for any $1 \leq i \leq n$. □

Lemma 3. *Let $p \in \mathcal{S}_K^{rat}(\Sigma)$. For any word $u \in res(p)$, $\sum_{w\in\Sigma^*} p(uw) \in K$. Moreover, the set $Res(p)$ generates the same subsemimodule of $K\langle\langle\Sigma\rangle\rangle$ as the set $\{\dot{u}p | u \in \Sigma^*\}$.*

Proof. Let $p \in \mathcal{S}_K^{rat}(\Sigma)$. For any word u, $\sum_{w\in\Sigma^*} p(uw) \in \mathbb{R}^+$ since p is a stochastic language. Suppose now that $K = \mathbb{Q}$ or $K = \mathbb{Q}^+$. The set $\{\dot{u}p | u \in \Sigma^*\}$ generates a finite vector subspace $\mathcal{P}$ of $\mathbb{Q}\langle\langle\Sigma\rangle\rangle$. Let $\{\dot{u}_1p,\ldots,\dot{u}_np\}$ be a finite subset of $\{\dot{u}p | u \in \Sigma^*\}$ which generates $\mathcal{P}$. Let $\sigma_i = \sum_{w\in\Sigma^*} \dot{u}_ip(w)$ for any $i = 1,\ldots,n$. From Lemma 2, each $\sigma_i \in \mathbb{Q}$. Now, for any $u \in \Sigma^*$, there exists $\alpha_1,\ldots,\alpha_n \in \mathbb{Q}$ such that $\dot{u}p = \sum_{i=1}^n \alpha_i\dot{u}_ip$. Therefore, $\sum_{w\in\Sigma^*} p(uw) = \sum_{i=1}^n \alpha_i\sigma_i \in \mathbb{Q}^+$.

So, for any K and any $u \in res(p)$, there exists an inversible element α_u of K such that $\dot{u}p = \alpha_u \cdot u^{-1}p$. In consequence, the set $Res(p)$ generates the same subsemimodule of $K\langle\langle\Sigma\rangle\rangle$ as the set $\{\dot{u}p | u \in \Sigma^*\}$. □

For any stochastic language p over K, let us denote by $[Res(p)]$ the subsemimodule of $K\langle\langle\Sigma\rangle\rangle$ generated by $Res(p)$ and let us call it the *residual subsemimodule* of p. Note that $[Res(p)]$ is stable.

Proposition 11. *Let $p \in \mathcal{S}_K^{rat}(\Sigma)$. For any word $u \in res(p)$, $u^{-1}p \in \mathcal{S}_K^{rat}(\Sigma)$.*

Proof. From Lemma 3, the residual stochastic languages $u^{-1}p$ belong to the same stable subsemimodules of $K\langle\langle\Sigma\rangle\rangle$ as p. Therefore, they are rational over K. □

3.3 Characterization of $\mathcal{S}_K^{rat}(\Sigma)$ in terms of stable subsemimodules

We show in this section that a series p over K is a rational stochastic language if and only if there exists a finite subset S in $\mathcal{S}(\Sigma)$ which generates a stable subsemimodule of $K\langle\langle\Sigma\rangle\rangle$ and such that $p \in conv_K(S)$.

The ≪ if part ≫ is easy to prove.

Proposition 12. *Let $p \in K\langle\langle\Sigma\rangle\rangle$. Suppose that there exists a finite subset S in $\mathcal{S}(\Sigma)$ which generates a stable subsemimodule of $K\langle\langle\Sigma\rangle\rangle$ and such that $p \in conv_K(S)$. Then $p \in \mathcal{S}_K^{rat}(\Sigma)$.*

Proof. Let $\{p_1, \ldots, p_n\}$ be a finite subset of $\mathcal{S}(\Sigma)$ which generates a stable subsemimodule of $K\langle\langle\Sigma\rangle\rangle$ and let $p = \sum_{i=1}^n \alpha_i p_i$ where $\alpha_i \geq 0$ for $i = 1, \ldots, n$ and $\sum_{i=1}^n \alpha_i = 1$. From Theorem 2, p is a rational series over K and p is a stochastic language since $p(w) = \sum_{i=1}^n \alpha_i p_i(w) \geq 0$ for any word w and $p(\Sigma^*) = \sum_{i=1}^n \alpha_i p_i(\Sigma^*) = 1$. □

The converse proposition is easy to prove when $K = \mathbb{Q}$ or $K = \mathbb{R}$. It is slightly more complicated when K is not a field.

Proposition 13. *Let $p \in \mathcal{S}_K^{rat}(\Sigma)$. Then there exists a finite subset S in $\mathcal{S}(\Sigma)$ which generates a stable subsemimodule of $K\langle\langle\Sigma\rangle\rangle$ and such that $p \in conv_K(S)$.*

Proof. Let $p \in \mathcal{S}_K^{rat}(\Sigma)$.

When $K = \mathbb{Q}$ or $K = \mathbb{R}$, K is a commutative field, $K\langle\langle\Sigma\rangle\rangle$ is a vector space and subsemimodules of $K\langle\langle\Sigma\rangle\rangle$ are vector subspaces of $K\langle\langle\Sigma\rangle\rangle$. From Lemma 3, the subspaces generated by $\{\dot{u}p | u \in \Sigma^*\}$ and $\{u^{-1}p | u \in \Sigma^*\}$ coincide. From Theorem 2, $\{u^{-1}p | u \in \Sigma^*\}$ generates a stable finite vector subspace $\mathcal{P}$ of $K\langle\langle\Sigma\rangle\rangle$. Let S be a finite subset of $\{u^{-1}p | u \in res(p)\}$ which contains p and generates $\mathcal{P}$. Clearly, $S \subseteq \mathcal{S}(\Sigma)$ and $p \in conv_K(S)$.

Let $K = \mathbb{Q}^+$ or $K = \mathbb{R}^+$. From Theorem 2, let $R = \{r_1, \ldots, r_n\}$ be a finite subset of $K\langle\langle\Sigma\rangle\rangle$ which generates a stable subsemimodule M containing p. We may suppose that $0 \notin R$ as R and $R \setminus \{0\}$ generate the same subsemimodule. Let $S = \{r \in R | \sum_{w\in\Sigma^*} r(w) < \infty\}$. First, let us show that S also generates a stable subsemimodule containing p. Let $T = R \setminus S$. Let $s \in S$ and let $u \in \Sigma^*$. As M is stable, we can write $\dot{u}s = \sum_{r\in R} \alpha_r^u r$, where the coefficients α_r^u belong to K. As $s \in S$, $\sum_{w\in\Sigma^*} \dot{u}s(w) < \infty$. Therefore, $r \in T \Rightarrow \alpha_r^u = 0$ and S generates a stable subsemimodule. In a similar way, we can write $p = \sum_{r\in R} \beta_r r$ and as p is a stochastic language, $r \in T \Rightarrow \alpha_r = 0$ and p belongs to the semimodule generated by S.

Now, let $S' = \{\left(\sum_{w\in\Sigma^*} s(w)\right)^{-1} \cdot s | s \in S\}$. Clearly, each element of S' is a stochastic language and an element of $K\langle\langle\Sigma\rangle\rangle$ (by using Lemma 2 when $K = \mathbb{Q}^+$). S' generates the same stable semimodule as S. We can write $p = \sum_{s\in S'} \beta_s s$, where the coefficients β_s belong to K. As p and each element of S' is a stochastic language, we have $\sum_{s\in S'} \beta_s = 1$ and hence, $p \in conv_K(S')$. □

Putting together the previous propositions, we obtain the following theorem:

Theorem 4. *Let $K \in \{\mathbb{R}, \mathbb{Q}, \mathbb{R}^+, \mathbb{Q}^+\}$. A series p over K is a rational stochastic language if and only if there exists a finite subset S in $\mathcal{S}(\Sigma)$ which generates a stable subsemimodule of $K\langle\langle\Sigma\rangle\rangle$ and such that $p \in conv_K(S)$.*

Proof. Apply Propositions 12 and 13. □

3.4 Subclasses of rational languages defined in terms of properties of their set of residual languages

Let p be a rational stochastic language over K. The set $Res(p)$ composed of the stochastic residual languages of p is *included* in a stable finitely generated subsemimodule of $K\langle\langle\Sigma\rangle\rangle$ but it may happen that the residual subsemimodule $[Res(p)]$ of p *is not* finitely generated. See Example 1 for instance. In the opposite, a stochastic language whose residual subsemimodule is finitely generated is rational. Therefore, two subclasses of $\mathcal{S}_K^{rat}(\Sigma)$ can be naturally defined:

- the set $\mathcal{S}_K^{fingen}(\Sigma)$ composed of rational stochastic languages over K whose residual subsemimodule is finitely generated;
- the set $\mathcal{S}_K^{fin}(\Sigma)$ composed of rational stochastic languages over K such that $Res(p)$ is finite.

Stochastic languages with finitely many residual languages. Every stochastic languages with finitely many residual languages can be described by using positive parameters only. In consequence, we obtain a Fatou-like property: every stochastic language with finitely many residual languages and which takes its values in $\mathbb{Q}$ is rational over $\mathbb{Q}^+$. Of course, for any K, there exist rational stochastic languages over K whose residual subsemimodule is finitely generated and which have not finitely many residual languages.

Proposition 14. *1.* $\mathcal{S}_{\mathbb{R}}^{fin}(\Sigma) = \mathcal{S}_{\mathbb{R}^+}^{fin}(\Sigma)$

2. $\mathcal{S}_{\mathbb{Q}}^{fin}(\Sigma) = \mathcal{S}_{\mathbb{Q}^+}^{fin}(\Sigma) = \mathcal{S}_{\mathbb{R}}^{fin}(\Sigma) \cap \mathbb{Q}^+\langle\langle\Sigma\rangle\rangle$.

3. For any $K \in \{\mathbb{R}, \mathbb{Q}, \mathbb{R}^+, \mathbb{Q}^+\}$, $\mathcal{S}_K^{fin}(\Sigma) \subsetneq \mathcal{S}_K^{fingen}(\Sigma)$.

Proof. 1. It is sufficient to show that $\mathcal{S}_{\mathbb{R}}^{fin}(\Sigma) \subseteq \mathcal{S}_{\mathbb{R}^+}^{fin}(\Sigma)$ in order to prove the first equality. Let $p \in \mathcal{S}_{\mathbb{R}}^{fin}(\Sigma)$ and let $Res(p) = \{u_1^{-1}p, \ldots, u_n^{-1}p\}$ be the set of residual languages of p. For any $u \in \Sigma^*$ and any $i \in \{1, \ldots, n\}$, there exists $j \in \{1, \ldots, n\}$ such that $\dot{u}u_i^{-1}p = u_i^{-1}p(u\Sigma^*)u_j^{-1}p$. Since $u_i^{-1}p(u\Sigma^*) \geq 0$, $Res(p)$ generates a stable subsemimodule of $\mathbb{R}^+\langle\langle\Sigma\rangle\rangle$. Since $p \in Res(p)$, $p \in \mathcal{S}_{\mathbb{R}^+}^{fin}(\Sigma)$ from Theorem 4.

2. The proof of the first equality goes in a similar way, with the complementary argument that $u_i^{-1}p(u\Sigma^*) \in \mathbb{Q}$ from Lemma 3.

   Now, let $p \in \mathcal{S}_{\mathbb{R}}^{fin}(\Sigma) \cap \mathbb{Q}^+\langle\langle\Sigma\rangle\rangle$. From Prop. 8, $p \in \mathcal{S}_{\mathbb{Q}}^{rat}(\Sigma)$. Therefore, $p \in \mathcal{S}_{\mathbb{Q}}^{fin}(\Sigma)$.

3. Consider the probabilistic automaton defined on Fig. 5. It defines a stochastic language p over $\mathbb{Q}^+$. Let us show that $p \in \mathcal{S}_{\mathbb{Q}^+}^{fingen}(\Sigma) \setminus \mathcal{S}_{\mathbb{Q}^+}^{fin}(\Sigma)$.

   First, let us show by induction on n that for any integer n, there exist $\alpha_n, \beta_n \in \mathbb{Q}^+$ such that $\overline{\dot{a^n}}p = \alpha_n p + \beta_n \dot{a}p$. This is true when $n = 0$: take $\alpha_0 = 1$ and $\beta_0 = 0$.

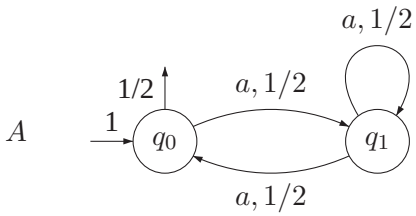

Fig. 5. The automaton A generates a stochastic language over $\mathbb{Q}^+$ whose residual subsemimodule is finitely generated but which has infinitely many residual languages.

Suppose that the relation holds for the integer n. For any word u, we have:

$$\begin{aligned}\dot{\overline{a^{n+1}}}p(u) &= \dot{\overline{a^n}}p(au)\\ &= \alpha_n p(au) + \beta_n \dot{a}p(au) \text{ by induction hypothesis}\\ &= \frac{\alpha_n}{2}\dot{a}p(u) + \beta_n\left(\frac{1}{2}p(u) + \frac{1}{2}\dot{a}p(u)\right) \text{ by remarking that } p = p_{q_0}\\ &\qquad\qquad \text{and } \dot{a}p = p_{q_1}.\end{aligned}$$

So we can take $\alpha_{n+1} = \beta_n/2$ and $\beta_{n+1} = (\alpha_n + \beta_n)/2$ which belong to $\mathbb{Q}^+$ from induction hypothesis. Therefore the module $[Res(p)]$ is finitely generated from Lemma 3: $p \in \mathcal{S}^{fingen}_{\mathbb{Q}^+}(\Sigma)$ and therefore, $p \in \mathcal{S}^{fingen}_{K}(\Sigma)$ for any $K \in \{\mathbb{R}, \mathbb{Q}, \mathbb{R}^+, \mathbb{Q}^+\}$.
Let $\gamma_n = (a^n)^{-1}p(\varepsilon)$. We have

$$\gamma_n = \frac{\alpha_n p(\varepsilon) + \beta_n \dot{a}p(\varepsilon)}{\alpha_n + \beta_n} = \frac{\alpha_n}{2(\alpha_n + \beta_n)}.$$

Check that γ_n satisfies the following induction relation:

$$\gamma_{n+1} = \frac{1 - 2\gamma_n}{4(1 - \gamma_n)}.$$

The sequence (γ_n) converges to the irrational number $(3 - \sqrt{5})/4$ and therefore, $\gamma_n = (a^n)^{-1}p(\varepsilon)$ takes an infinite number of values, which implies that p has infinitely many residual languages. □

Stochastic languages whose residual subsemimodule is finitely generated . When K is a field, every rational stochastic language is finitely generated. This property is no longer true when $K \in \{\mathbb{R}^+, \mathbb{Q}^+\}$. In consequence, some stochastic languages whose residual subsemimodule is finitely generated cannot be generated by using only positive parameters.

We prove also a Fatou-like property: every stochastic language over $\mathbb{R}^+$ whose residual subsemimodule is finitely generated and which takes its values in $\mathbb{Q}$ is rational over $\mathbb{Q}^+$. But we first need the following technical lemmas.

Lemma 4. *Let $k, n \in \mathbb{N}$ and let $\alpha_i, \beta_i^j \in \mathbb{Q}$for $1 \leq i \leq n$ and $1 \leq j \leq k$. Consider the variables $x_1, \ldots, x_k$ and the system (S) composed of the n following inequations*

$$\alpha_i + \sum_{j=1}^{k} x_j \beta_i^j \geq 0$$

for $i = 1, \ldots, n$. If (S) has a solution, then it has also a solution which satisfies

$$\alpha_i + \sum_{j=1}^{k} x_j \beta_i^j \in \mathbb{Q}^+$$

for $i = 1, \ldots, n$.

Proof. By induction on n.

- Let $n = 1$. Let $\mu_1, \ldots, \mu_k$ be such that $\alpha_1 + \sum_{j=1}^{k} \mu_j \beta_1^j \geq 0$. If $\alpha_1 + \sum_{j=1}^{k} \mu_j \beta_1^j = 0$, we are done. If $\alpha_1 + \sum_{j=1}^{k} \mu_j \beta_1^j > 0$, there exists $\mu'_1, \ldots, \mu'_k \in \mathbb{Q}$ such that $\alpha_1 + \sum_{j=1}^{k} \mu'_j \beta_1^j > 0$ since $\mathbb{Q}$ is dense in $\mathbb{R}$ and since $\alpha_1 + \sum_{j=1}^{k} \mu_j \beta_1^j$ is a continuous expression of the μ_i.
- Let $n > 1$ and let $\mu_1, \ldots, \mu_k$ be such that $\alpha_i + \sum_{j=1}^{k} \mu_j \beta_i^j \geq 0$ for any $1 \leq i \leq n$. If $\alpha_i + \sum_{j=1}^{k} \mu_j \beta_i^j > 0$ for any integer i, then there exists $\mu'_1, \ldots, \mu'_k \in \mathbb{Q}$ such that $\alpha_i + \sum_{j=1}^{k} \mu'_j \beta_i^j > 0$ for any i, by using the same argument as previously. Otherwise, there exists at least an integer i such that $\alpha_i + \sum_{j=1}^{k} \mu_j \beta_i^j = 0$.
  - If each $\beta_i^j = 0$, then α_i is also null and this equation can be ruled out from the system without modifying its solutions. In this case, the induction hypothesis can be directly applied.
  - If there exists j such that $\beta_i^j \neq 0$, then μ_j can be expressed as a function of the other μ_i: $\mu_j = -(\alpha_i + \sum_{l \neq j} \mu_l \beta_i^l)/\beta_i^j$, x_j can be replaced with $-(\alpha_i + \sum_{l \neq j} x_l \beta_i^l)/\beta_i^j$ in all the other inequations and the induction hypothesis can be applied.

□

Lemma 5. *Let $r_0, r_1, \ldots, r_n \in \mathbb{Q}\langle\langle \Sigma \rangle\rangle$ and let $\alpha_1, \ldots, \alpha_n \in \mathbb{Q}$, $\beta_1, \ldots, \beta_n \in \mathbb{R}^+$ be such that*

$$r_0 = \sum_{i=1}^{n} \alpha_i r_i = \sum_{i=1}^{n} \beta_i r_i.$$

Then, there exists $\gamma_1, \ldots, \gamma_n \in \mathbb{Q}^+$ such that

$$r_0 = \sum_{i=1}^{n} \gamma_i r_i.$$

Proof. The set of parameters $\{(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n | \sum_{i=1}^{n} \lambda_i r_i = 0\}$ is a vector subspace of $\mathbb{R}^n$. Since the series $r_1, \ldots, r_n$ take their values in $\mathbb{Q}$, there exist k vectors $(t_1^1, \ldots, t_n^1), \ldots, (t_1^k, \ldots, t_n^k) \in \mathbb{Q}^n$, with $k \leq n$, such that for any $(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$,

$$\sum_{i=1}^{n} \lambda_i r_i = 0 \text{ iff } \exists \mu_1, \ldots, \mu_k \in \mathbb{R} \text{ s.t. } \lambda_i = \sum_{j=1}^{k} \mu_j t_i^j \text{ for any } i = 1, \ldots, n.$$

Hence, for any $(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$,

$$r_0 = \sum_{i=1}^{n} \lambda_i r_i \text{ iff } \exists \mu_1, \ldots, \mu_k \in \mathbb{R} \text{ s.t. } \lambda_i = \alpha_i + \sum_{j=1}^{k} \mu_j t_i^j \text{ for any } i = 1, \ldots, n.$$

In particular, there exist $\mu_1, \ldots, \mu_k$ such that $\beta_i = \alpha_i + \sum_{j=1}^{k} \mu_j t_i^j \geq 0$ for any $i = 1, \ldots, n$.

Consider the system composed of the n inequations $\alpha_i + \sum_{j=1}^{k} x_j t_i^j \geq 0$ for $i = 1, \ldots, n$. It has a solution and from the previous Lemma, it has also a solution $(\mu_1, \ldots, \mu_k)$ which satisfies $\alpha_i + \sum_{j=1}^{k} \mu_j t_i^j \in \mathbb{Q}^+$ for $i = 1, \ldots, n$. □

Proposition 15. *1. When $K \in \{\mathbb{R}, \mathbb{Q}\}$, $\mathcal{S}_K^{fingen}(\Sigma) = \mathcal{S}_K^{rat}(\Sigma)$.*
2. When $K \in \{\mathbb{R}^+, \mathbb{Q}^+\}$, $\mathcal{S}_K^{fingen}(\Sigma) \subsetneq \mathcal{S}_K^{rat}(\Sigma)$.
3. $\mathcal{S}_{\mathbb{Q}^+}^{fingen}(\Sigma) = \mathcal{S}_{\mathbb{R}^+}^{fingen}(\Sigma) \cap \mathbb{Q}^+\langle\langle\Sigma\rangle\rangle$.

Proof. 1. When $K \in \{\mathbb{R}, \mathbb{Q}\}$, K is a commutative field. As a consequence, any vector subspace of a finitely generated vector subspace of $K\langle\langle\Sigma\rangle\rangle$ is finitely generated itself. Therefore, for any $p \in \mathcal{S}_K^{rat}(\Sigma)$, the residual subsemimodule of p is finitely generated.

2. Example 1 describes a rational stochastic language whose residual subsemimodule is not finitely generated.

3. Let $p \in \mathcal{S}_{\mathbb{R}^+}^{fingen}(\Sigma) \cap \mathbb{Q}^+\langle\langle\Sigma\rangle\rangle$. Let $S = \{r_1, \ldots, r_n\} \subseteq Res(p)$ be a finite subset which generates the same subsemimodule as $Res(p)$ in $\mathbb{R}^+\langle\langle\Sigma\rangle\rangle$. From Prop. 8, $p \in \mathcal{S}_{\mathbb{Q}}^{rat}(\Sigma)$ and from Prop. 11, each $r_i \in \mathcal{S}_{\mathbb{Q}}^{rat}(\Sigma)$. S also generates the same subsemimodule as $Res(p)$ in $\mathbb{Q}\langle\langle\Sigma\rangle\rangle$. From Lemma 5, for any word u and any index i, there exists $\gamma_1^{i,u}, \ldots, \gamma_n^{i,u} \in \mathbb{Q}^+$ such that $\dot{u}r_i = \sum_{j=1}^{n} \gamma_j^{i,u} r_j$. Therefore, S generates a stable subsemimodule of $\mathbb{Q}^+\langle\langle\Sigma\rangle\rangle$. Also from Lemma 5, there exists $\gamma_1, \ldots, \gamma_n \in \mathbb{Q}^+$ such that $p = \sum_{i=1}^{n} \gamma_i r_i$. Therefore, $p \in conv_{\mathbb{Q}^+}(S)$ and $p \in \mathcal{S}_{\mathbb{Q}^+}^{fingen}(\Sigma)$.

□

Remark that $\mathcal{S}_{\mathbb{Q}^+}^{fingen}(\Sigma) \subsetneq \mathcal{S}_{\mathbb{R}}^{fingen}(\Sigma) \cap \mathbb{Q}^+\langle\langle\Sigma\rangle\rangle$ since $\mathcal{S}_{\mathbb{Q}^+}^{fingen}(\Sigma) \subsetneq \mathcal{S}_{\mathbb{Q}}^{rat}(\Sigma) = \mathcal{S}_{\mathbb{R}}^{rat}(\Sigma) \cap \mathbb{Q}^+\langle\langle\Sigma\rangle\rangle = \mathcal{S}_{\mathbb{R}}^{fingen}(\Sigma) \cap \mathbb{Q}^+\langle\langle\Sigma\rangle\rangle$

Finaly, we show that when K is positive, finitely generated stochastic languages over K have a unique normal representation in terms of stable subbsemimodules generated by residual languages which is minimal with respect to inclusion.

Proposition 16. *Let $K = \mathbb{Q}^+$ or $K = \mathbb{R}^+$ and let $p \in \mathcal{S}_K^{fingen}(\Sigma)$. Then, there exists a unique finite subset $R \subseteq Res(p)$ which generates a stable subsemimodule of $K\langle\langle\Sigma\rangle\rangle$, such that $p \in conv_K(R)$ and which is minimal for inclusion.*

Proof. Let $K = \mathbb{Q}^+$ or $K = \mathbb{R}^+$ and let $p \in \mathcal{S}_K^{fingen}(\Sigma)$. Let $R = \{r_1, \ldots, r_n\}$ and $S = \{s_1, \ldots, s_m\}$ be two minimal subsets of $Res(p)$ generating $[Res(p)]$. Let $r_{i_0} \in R$. We are to prove that $r_0 \in S$.

There exist $\alpha_{i_0}^1, \ldots, \alpha_{i_0}^n \in K$ such that $r_{i_0} = \sum_{i=1}^{m} \alpha_{i_0}^i s_i$.

There exist $\beta_i^j \in K$ for any $1 \leq i, j \leq n$ such that $s_i = \sum_{j=1}^{n} \beta_i^j r_j$ for any $1 \leq i \leq m$.

Therefore,

$$r_{i_0} = \sum_{i=1}^{m} \alpha_{i_0}^i \sum_{j=1}^{n} \beta_i^j r_j = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} \alpha_{i_0}^i \beta_i^j \right) r_j.$$

If $\sum_{i=1}^{m} \alpha_{i_0}^i \beta_i^{i_0} < 1$, then we could express r_{i_0} as a convex combination of the other r_i and R would not be minimal for inclusion. Therefore, $\sum_{i=1}^{m} \alpha_{i_0}^i \beta_i^{i_0} = 1$.

Since $\sum_{i=1}^{m} \alpha_{i_0}^i = 1$ and each $\beta_i^j \in [0,1]$, for any index i such that $\alpha_{i_0}^i \neq 0$, we must have $\beta_i^{i_0} = 1$. Therefore, for any index i such that $\alpha_{i_0}^i \neq 0$, we must have $s_i = r_{i_0}$. As such an index must exist, $r_{i_0} \in S$.

Since no condition has been put on r_{i_0}, then $R \subseteq S$ and finally, $R = S$. □

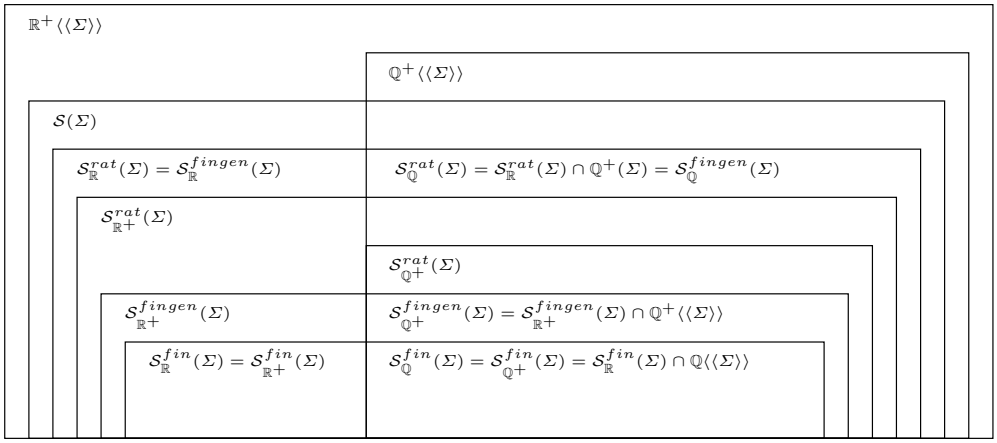

Fig. 6. Inclusion relations between classes of classes of rational stochastic languages, including $\mathcal{S}_K^{fingen}(\Sigma)$ and $\mathcal{S}_K^{fin}(\Sigma)$.

4 Multiplicity automata and rational stochastic languages.

In the previous Sections, we have defined several classes of rational stochastic languages over $K \in \{\mathbb{R}, \mathbb{Q}, \mathbb{R}^+, \mathbb{Q}^+\}$. In this section, we study the representation of these classes by means of multiplicity automata: given a subclass $\mathcal{C}$ of rational stochastic languages over K, is there a subset of K-multiplicity automata both simple to identify and sufficient to generate the elements of $\mathcal{C}$? The first result we prove is negative: it is undecidable whether a given multiplicity automaton over $\mathbb{Q}$ generates a stochastic language. Moreover, there exist no recursively enumerable subset of multiplicity automata over $\mathbb{Q}$ sufficient to generate $\mathcal{S}_{\mathbb{Q}}^{rat}(\Sigma)$. This result implies that no classes of multiplicity automata can efficiently represent the class of rational stochastic languages over $\mathbb{Q}$ or $\mathbb{R}$. In the other hand, we show that the class of *K-probabilistic automata* represents $\mathcal{S}_K^{rat}(\Sigma)$ when $K \in \{\mathbb{R}^+, \mathbb{Q}^+\}$. Clearly, it can be decided efficiently whether a given multiplicity automaton is a probabilistic automaton. We show also that the class of *K-probabilistic residual automata* represents the class $\mathcal{S}_K^{fingen}(\Sigma)$ for any

$K \in \{\mathbb{R}, \mathbb{R}^+, \mathbb{Q}, \mathbb{Q}^+\}$. We do not know whether the class of *probabilistic residual automata* is decidable. However, we show that it contains a subclass which is decidable and sufficient to generate $\mathcal{S}_K^{fingen}(\Sigma)$. Nevertheless, we show that deciding whether a given MA is in this subclass is a PSPACE-complete. Finally, the class of *probabilistic deterministic automata* over $\mathbb{R}^+$ (resp. $\mathbb{Q}^+$), which is clearly decidable, represents the class $\mathcal{S}_K^{fin}(\Sigma)$ when $K \in \{\mathbb{R}, \mathbb{R}^+\}$ (resp. $K \in \{\mathbb{Q}, \mathbb{Q}^+\}$).

To our knowledge, the decidability of the following problems is still open:

- decide whether a given multiplicity automaton is equivalent to a probabilistic automaton, or a probabilistic residual automaton or a probabilistic deterministic automaton;
- decide whether a given probabilistic automaton is equivalent to a probabilistic residual automaton or a probabilistic deterministic automaton;
- decide whether a given probabilistic residual automaton is equivalent to a probabilistic deterministic automaton.

4.1 The class of MA which generate stochastic languages is undecidable

A MA A generates a stochastic language p_A if and only if

- $\forall w \in \Sigma^*, p_A(w) \geq 0$ and,
- $\sum_{w \in \Sigma^*} p_A(w) = 1$.

We first show that the second condition can be checked within polynomial time. We need the following result:

Lemma 6. *[Gan66,BT00] Let M be a square matrix with coefficients in $\mathbb{Q}$. It is decidable within polynomial time whether M^k converges to 0 when k tends to infinity.*

Proof. (Sketch) First, M^k converges to 0 when k tends to infinity if and only if the spectral radius $\rho(M)$ of M, i.e. the maximum of the magnitudes of its eigenvalues, satisfies $\rho(M) < 1$.

Then, M satisfies $\rho(M) < 1$ iff the Lyapunov equation

$$MPM^t = P$$

has a positive-definite solution. In that case the solution is unique. Since the Lyapunov equation is linear in the unknown entries of P, we can compute a a solution P in polynomial time, or decide it does not exist. To check that P is positive definite, it is sufficient to compute the determinants of the principal minors of P and check that they are all positive. □

Proposition 17. *Let A be an MA over $\mathbb{Q}$. It is decidable within polynomial time whether the sum $\sum_k P_A(\Sigma^k)$ converges. If the sum $P_A(\Sigma^*) = \sum_k P_A(\Sigma^k)$ converges, it can be computed within polynomial time.*

Proof. Let $A = \langle \Sigma, Q, \varphi, \iota, \tau \rangle$ where $Q = \{q_1, \ldots, q_n\}$ and let M be the square matrix defined by $M[i,j] = [\varphi(q_i, \Sigma, q_j)]_{1 \leq i,j \leq n}$. We have $P_A(\Sigma^k) = \iota_A M^k \tau_A$ where $\iota_A = (\iota(q_1), \ldots, \iota(q_n))$ and $\tau_A = (\tau(q_1), \ldots, \tau(q_n))^t$.

Let E be the subspace of $\mathbb{R}^n$ spanned by $\{M^k\tau_A | k \in \mathbb{N}\}$ and let F be a complementary subspace of E in $\mathbb{R}^n$. Let $H = \{u \in E | \forall k \in \mathbb{N},\ \iota_A M^k u = 0\}$. Clearly, E and H are stable under M. Let G be a complementary subspace of H in E. For any $u \in \mathbb{R}^n$, there exists a unique decomposition of the form $u = u_F + u_G + u_H$ where $u_F \in F, u_G \in G$ and $u_H \in H$. Let p_F, p_H and p_G be the projections on F, G and H defined by $p_F(u) = u_F$, $p_G(u) = u_g$ and $p_H(u) = u_H$. Let P_F, P_H and P_G be the corresponding matrices.

First note that for any integer $k \geq 1$ and any $u \in E$, we have $P_G M^k P_G u = (P_G M P_G)^k u$. This is clear when $k = 1$. We have

$$\begin{aligned} P_G M^{k+1} P_G u &= P_G M^k (M P_G u) \\ &= P_G M^k [P_H M P_G u + P_G M P_G u] \text{ since } M P_G u \in E \\ &= P_G M^k P_G [P_G M P_G u] \text{ since } \forall v \in H, Mv \in H \text{ and } P_G(v) = 0 \\ &= (P_G M P_G)^{k+1} u \text{ from induction hypothesis.} \end{aligned}$$

Note also that for any integer k and any $u \in E$,

$$\begin{aligned} \iota_A M^k u &= \iota_A M^k (P_G u + P_H u) \text{ since } u \in E \\ &= \iota_A M^k P_G u \text{ since } \forall v \in H, Mv \in H \text{ and } \iota_A v = 0 \\ &= \iota_A (P_G M^k P_G u + P_H M^k P_G u) \text{ since } M^k P_G u \in E \\ &= \iota_A P_G M^k P_G u \text{ since } \forall v \in H, \iota_A v = 0 \\ &= \iota_A (P_G M P_G)^k u. \end{aligned}$$

We show now that $\sum_{k\in\mathbb{N}} \iota_A M^k \tau_A$ is convergent iff $\lim_{k\to\infty}(P_G M P_G)^k = 0$.

– Suppose that $\lim_{k\to\infty}(P_G M P_G)^k = 0$. Then $Id - P_G M P_G$ is inversible and $\sum_{k\in\mathbb{N}}(P_G M P_G)^k$ converges to $(Id - P_G M P_G)^{-1}$. Therefore, $\sum_{k\in\mathbb{N}} \iota_A M^k \tau_A$ converges to $\iota_A (Id - P_G M P_G)^{-1} \tau_A$.
– Suppose now that $\sum_{k\in\mathbb{N}} \iota_A M^k \tau_A$ is convergent.
There exists $\lambda > 0$ such that for all $u \in G$, there exists $n \in \mathbb{N}$ such that $|\iota_A M^n u| \geq \lambda ||u||$. Otherwise, there would exist a sequence u_k of elements of G such that for all integer n, $|\iota_A M^n (u_k)| < ||u_k||/k$. Let $v_k = u_k/||u_k||$ and let $v_{\sigma(k)}$ a subsequence which converges to v. Check that we should have $||v|| = 1$, $v \in G$ and $\iota_A M^n v = 0$ for any integer n, which is impossible since $v \neq 0$.
Let λ satisfying this property. For any integers m and k, there exists n_k such that

$$|\iota_A M^{n_k} (P_G M^k P_G)(M^m \tau_A)| \geq \lambda ||(P_G M^k P_G)(M^m \tau_A)|| = \lambda ||(P_G M P_G)^k (M^m \tau_A)||.$$

We have also

$$\begin{aligned} \iota_A M^{n_k} (P_G M^k P_G)(M^m \tau_A) &= \iota_A (P_G M P_G)^{n_k} (P_G M^k P_G)(M^m \tau_A) \\ &= \iota_A (P_G M P_G)^{n_k + k} (M^m \tau_A) \\ &= \iota_A M^{n_k + k} (M^m \tau_A) \\ &= \iota_A M^{n_k + k + m} \tau_A. \end{aligned}$$

If we suppose that $\iota_A M^k \tau_A \to 0$ when $k \to \infty$, we must have $|(P_G M^k P_G)(M^m \tau_A)|| \to 0$ when $k \to \infty$ for any integer m. As $\{M^m \tau_A\}$ generates E, $P_G M^k P_G$ converges to 0.

To sum up, $\sum_k P_A(\Sigma^k)$ is bounded iff $(P_G M P_G)^k$ converges to 0, which is a polynomially decidable problem (Lemma 6).

When the sum $\sum_k P_A(\Sigma^k)$ converges, it is equal to $\iota_A(Id - P_G M P_G)^{-1}\tau_A$ which can be computed within polynomial time. □

Example 2. Consider the MA A'' described on Fig. 3. We have

$$\iota_{A''} = (1,0), \tau_{A''} = (1/4, 1/4)^t \text{ and } M = \begin{pmatrix} \frac{3}{4} & 0 \\ 0 & \frac{3}{4} \end{pmatrix}$$

We have $M\tau_{A''} = 3/4\tau_{A''}$ and therefore, E is the vector space spanned by $\tau_{A''}$. Let F be the complementary space of E spanned by the vector $(1,-1)^t$; we have

$$H = \{0\}, G = E, P_G = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \text{ and } 1 - P_G M P_G = \frac{1}{8}\begin{pmatrix} 5 & -3 \\ -3 & 5 \end{pmatrix}$$

Check that the inverse of $1 - P_G M P_G$ is equal to

$$\frac{1}{2}\begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$$

and that $\iota_A(Id - P_G M P_G)^{-1}\tau_A = 1$.

We prove now that it is undecidable whether a multiplicity over $\mathbb{Q}$ generates a stochastic language. In order to prove this result, we use a reduction to a decision problem about *acceptor PAs*.

An MA $\langle \Sigma, Q, \varphi, \iota, \tau \rangle$ is an *acceptor PA* if

- φ, ι and τ are non negative functions,
- $\sum_{q \in Q} \iota(q) = 1$,
- $\forall q \in Q, \forall x \in \Sigma, \sum_{r \in Q} \varphi(q,x,r) = 1$
- there exists a unique terminal state t and $\tau(t) = 1$.

Blondel and Canterini have shown that given an acceptor PA A over $\mathbb{Q}$ and $\lambda \in \mathbb{Q}$, it is undecidable whether there exists a word w such that $P_A(w) < \lambda$ ([BC03]).

Theorem 5. *It is undecidable whether an MA over $\mathbb{Q}$ generates a stochastic language.*

Proof. For any rational series r over Σ, let us denote by $\overline{r}$ the rational series defined by

$$\overline{r} = \sum_{w \in \Sigma^*} \frac{r(w)}{(|\Sigma|+1)^{|w|+1}}.$$

Let $A = \langle \Sigma, Q, \varphi, \iota, \tau \rangle$ be an acceptor PA over $\mathbb{Q}$ and let $\lambda \in \mathbb{Q}$. Let $B = \langle \Sigma, Q, \varphi_B, \iota, \tau_B \rangle$ be the MA defined by $\varphi_B(q,x,q') = \varphi(q,x,q')/(|\Sigma|+1)$ and $\tau_B(q) = \tau(q)/(|\Sigma|+1)$ for any states $q, q' \in Q$ and any $x \in \Sigma$. Remark that B is semi PA and that $r_B = \overline{r_A}$.

The sum $s = \sum_{w \in \Sigma^*} r_B(w)$ is bounded by 1 from Prop. 2 and can be computed within polynomial time by using the Prop. 17. Let c_λ be the series defined by $c_\lambda(w) = \lambda$ for any word $w \in \Sigma^*$.

– If $s < \lambda$, then there must exists a word w such that $P_A(w) < \lambda$ since

$$\sum_{w \in \Sigma^*} \frac{\lambda}{(|\Sigma|+1)^{|w|+1}} = \lambda.$$

– If $s = \lambda$, the rational series $1 + \overline{r_A - c_\lambda}$ is a stochastic language iff $r_A(w) \geq \lambda$ for any word w.
– If $s > \lambda$, the rational series $\frac{1}{s-\lambda} \cdot \overline{r_A - c_\lambda}$ is a stochastic language iff $r_A(w) \geq \lambda$ for any word w.

Since in the two last cases, a multiplicity automaton which generates $1+\overline{r_A - c_\lambda}$ (resp. $\frac{1}{s-\lambda} \cdot \overline{r_A - c_\lambda}$) can easily be derived from A, an algorithm able to decide whether an MA generates a stochastic language could be used to solve the decision problem on PA acceptors. □

A reduction to the following undecidable problem could have also been used: it is undecidable whether a rational series over $\mathbb{Z}$ takes a negative value [SS78].

The set of multiplicity automata over $\mathbb{Q}$ which generate stochastic languages is not only not recursive: it contains no recursively enumerable set able to generate $\mathcal{S}^{rat}_{\mathbb{Q}}(\Sigma)$.

Theorem 6. *No recursively enumerable set of multiplicity automata over $\mathbb{Q}$ exactly generates $\mathcal{S}^{rat}_{\mathbb{Q}}(\Sigma)$.*

Proof. From Prop. 17, the set $\mathcal{A}$ composed of the multiplicity automata A over $\mathbb{Q}$ which satisfy $P_A(\Sigma^*) = 1$ is recursively enumerable.

The subset $\mathcal{B}$ composed of the elements of $\mathcal{A}$ which satisfy

$$\exists\, w \in \Sigma^* P_A(w) < 0$$

is recursively enumerable.

Suppose that there exists a recursive enumeration $R_0, \ldots, R_n, \ldots$ of multiplicity automata over $\mathbb{Q}$ sufficient to generate $\mathcal{S}^{rat}_{\mathbb{Q}}(\Sigma)$ and let $w_0, \ldots, w_n, \ldots$ be an enumeration of Σ^*.

Consider the following algorithm:

```
Input: a multiplicity automaton A over ℚ
If p_A(Σ*) = 1 then
    For i ≥ 0 do
        If p_A(w_i) < 0 then output NO; exit; EndIf
        If A is equivalent to R_i then output YES; exit; EndIf
    EndFor
Else
    output NO; exit
EndIf
```

Since the equality $\sum_{w \in \Sigma^*} P_A(w) = 1$ and the equivalence of two multiplicity automata can be decided, this algorithm would end on any input and decide whether A generates a stochastic language. Therefore, the enumeration $R_0, \ldots, R_n, \ldots$ cannot exist. □

4.2 Probabilistic automata

So, $\mathcal{S}^{rat}_{\mathbb{Q}}(\Sigma)$ and $\mathcal{S}^{rat}_{\mathbb{R}}(\Sigma)$ cannot be identified by any efficient subclass of multiplicity automata. In the other hand, $\mathcal{S}^{rat}_{\mathbb{Q}^+}(\Sigma)$ and $\mathcal{S}^{rat}_{\mathbb{R}^+}(\Sigma)$ can be described by probabilistic automata which form an easily identifiable subclass of multiplicity automata.

Proposition 18. *Let $K \in \{\mathbb{R}^+, \mathbb{Q}^+\}$ and let $p \in K\langle\langle\Sigma\rangle\rangle$. Then, p is a stochastic language over K iff there exists a K-probabilistic automaton A such that $p = r_A$.*

Proof. The only thing to prove is that if $p \in \mathcal{S}^{rat}_K(\Sigma)$ then there exists a K-probabilistic automaton A such that $p = r_A$.

From Theorem 4, there exist a finite subset S of $\mathcal{S}^{rat}_K(\Sigma)$ which generates a stable subsemimodule of $K\langle\langle\Sigma\rangle\rangle$ and such that $p \in conv_K(S)$. Suppose that S is minimal for inclusion. For any $s, s' \in S$ and any $x \in \Sigma$, let α_s and $\alpha^x_{s,s'} \in K$ such that $p = \sum_{s\in S} \alpha_s s$ and $\dot{x}s = \sum_{s'\in S} \alpha^x_{s,s'} s'$.

Let $A = \langle \Sigma, S, \varphi, \iota, \tau\rangle$ be the MA defined by:

- $\iota(s) = \alpha_s$,
- $\tau(s) = s(\varepsilon)$,
- $\varphi(s, x, s') = \alpha^x_{s,s'}$

for any $s, s' \in S$ and any $x \in \Sigma$. From Claims 1 and 2, $p = r_A$.

Since $S \subseteq \mathcal{S}^{rat}_K(\Sigma)$, every state of A is co-accessible and since S is minimal, every state of A is accessible. Therefore, A is trimmed.

Note that $\sum_{s\in S} \iota(s) = \sum_{s\in S} \alpha_s = 1$ since elements of $\{p\} \cup S$ are stochastic languages. For any $s \in S$,

$$\begin{aligned}
\tau(s) + \sum_{s'\in S, x\in\Sigma} \varphi(s, x, s') &= s(\varepsilon) + \sum_{s'\in S, x\in\Sigma} \alpha^x_{s,s'} \\
&= s(\varepsilon) + \sum_{x\in\Sigma} \dot{x}s(\Sigma^*) \\
&= s(\varepsilon) + \sum_{x\in\Sigma} s(x\Sigma^*) \\
&= 1.
\end{aligned}$$

Then, A is a PA. □

4.3 Probabilistic residual automata

For any $K \in \{\mathbb{R}^+, \mathbb{Q}^+\}$, the class $\mathcal{S}^{fingen}_K(\Sigma)$ can be described by probabilistic residual automata.

Proposition 19. *Let $K \in \{\mathbb{R}^+, \mathbb{Q}^+\}$ and let $p \in K\langle\langle\Sigma\rangle\rangle$. Then, p is a stochastic language over K whose residual subsemimodule is finitely generated iff there exists a K-probabilistic residual automaton A such that $p = r_A$.*

Proof. – Let $p \in \mathcal{S}^{fingen}_K(\Sigma)$ and let $w_1, \ldots, w_n \in res(p)$ be such that $S = \{w_1^{-1}p, \ldots, w_n^{-1}p\}$ generates $[Res(p)]$. Let A be the MA associated with S as in the proof of Prop. 18. Check that A is a PRA which generates p.

– Let $A\langle\Sigma, Q, \varphi, \iota, \tau\rangle$ be a PRA which generates p and for any $q \in Q$, let $w_q \in \Sigma^*$ be such that $r_{A,q} = w_q^{-1}p$. From Claim 3, $\{w_q^{-1}p | q \in Q\}$ generates a stable subsemimodule M which contains p. Check that $[Res(p)] = M$.

□

Remark that from Prop. 16, there exists a unique minimal subset S of $Res(p)$ which generates $[Res(p)]$. A PRA based on this set has a minimal number of states.

We do not know whether the class of PRA is decidable. However, we show that the class of $\mathbb{R}^+$-*reduced PRA* is decidable. Since a reduced PRA is a PRA, any PRA is equivalent to a reduced PRA and therefore, this class is sufficient to generate $\mathcal{S}_K^{fingen}(\Sigma)$.

Let A be a PA and let $\langle\Sigma, Q, \delta, Q_I, Q_T\rangle$ be the support of A. If for any state $q \in Q$, there exists a word w_q such that $\delta(Q_I, w_q) = \{q\}$, then A is a PRA since $w_q^{-1}r_A = r_{A,q}$. The converse is true when A is reduced.

Proposition 20. *Let A be a $\mathbb{R}^+$-reduced PA and let $\langle\Sigma, Q, \delta, Q_I, Q_T\rangle$ be the support of A. Then, A is a PRA if and only if for any state $q \in Q$, there exists a word w such that $\delta(Q_I, w) = \{q\}$.*

Proof. Suppose that A is a PRA. Let $q \in Q$ and w be a word such that $w_q^{-1}r_A = r_{A,q}$. Let $Q_w = \delta(Q_I, w)$. There exist $(\alpha_{q'})_{q' \in Q_w}$ such that $w^{-1}r_A = \sum_{q' \in Q_w} \alpha_{q'} r_{A,q'}$. Since $q \in Q_w$, $(1 - \alpha_q)r_{A,q} = \sum_{q' \in Q_w, q' \neq q} \alpha_{q'} r_{A,q'}$. Since A is $\mathbb{R}^+$-reduced, we must have $\alpha_q = 1$ and therefore, $Q_w = \{q\}$. □

Corollary 1. *It can be decided whether a $\mathbb{R}^+$-reduced MA is a PRA.*

Proof. It can easily be decided whether an MA is a PA. Then, the power set construction can be used to check whether any state can be uniquely reached by some word.

□

From Prop. 6, it can efficiently be decided whether an MA is $\mathbb{R}^+$-reduced PA. But unfortunately, no *efficient* decision procedure exist to decide whether it is an $\mathbb{R}^+$-reduced PRA: the decision problem is PSPACE-complete.

Proposition 21. *Deciding whether a $\mathbb{R}^+$-reduced PA is a PRA is PSPACE-complete.*

Proof. We prove the proposition by reduction of the following PSPACE-complete problem: given n DFA $A^1, \ldots, A^n$ over Σ, let L_i be the language recognized by A^i for $1 \leq i \leq n$, deciding whether $\cup_{i=1}^n L_i = \Sigma^*$ is PSPACE-complete.

Let $A^i = \langle\Sigma, Q^i, \{q_0^i\}, Q_T^i, \delta^i\rangle$ for $1 \leq i \leq n$ where $i \neq j$ implies that $Q^i \cap Q^j = \emptyset$. We may suppose that $L_i \neq \emptyset$ for $1 \leq i \leq n$. Consider 3 new states q_0, q_1, q_f, $n+1$ new letters $x_1, \ldots, x_n, \lambda$. Let $A = \langle\Sigma_A, Q_A, Q_I, Q_T, \delta\rangle$ be an NFA defined by:

– $\Sigma_A = \Sigma \cup \{x_1, \ldots, x_n, \lambda\}$
– $Q_A = \cup_{i=1}^n Q^i \cup \{q_0, q_1, q_f\}$,
– $Q_I = \{q_0, q_0^1, \ldots, q_0^n\}$,
– $Q_T = \{q_1, q_f\}$,
– for any $1 \leq i, j \leq n$, any $q \in Q^i$ and any $x \in \Sigma$,
  - $\delta(q, x) = \delta^i(q, x)$,
  - $\delta(q, x_j) = \{q_0^i\}$ if $i = j$ and $\emptyset$ otherwise,
  - $\delta(q, \lambda) = \{q_f\}$ if $q \in Q_T^i$ and $\emptyset$ otherwise,

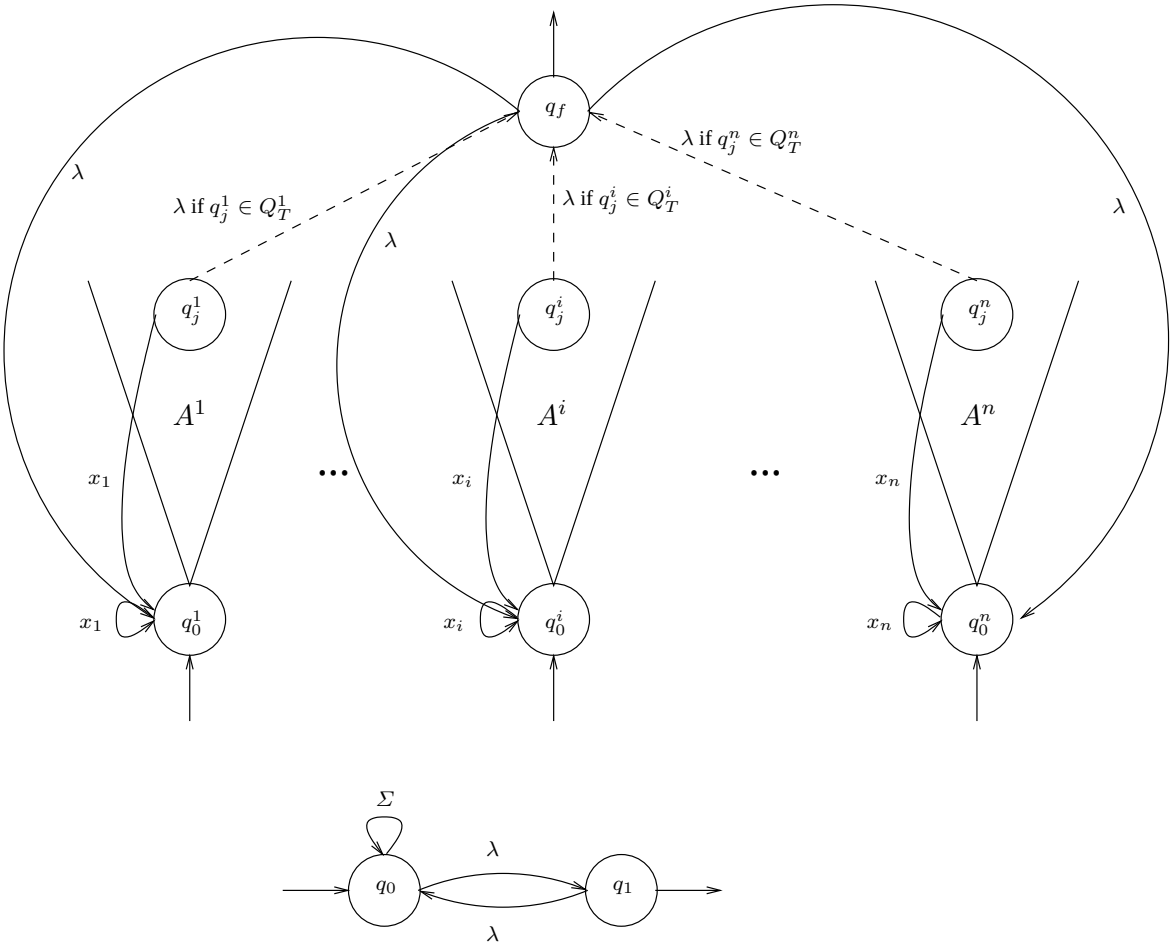

Fig. 7. The union of the languages recognized by the automata A_i is different from Σ^* if and only if this automaton is the support of a $\mathbb{R}^+$-reduced PRA.

– for any $x \in \Sigma$, $\delta(q_0, x) = \{q_0\}$, $\delta(q_1, x) = \emptyset$ and $\delta(q_f, x) = \emptyset$,
– $\delta(q_0, \lambda) = \{q_1\}$, $\delta(q_1, \lambda) = \{q_0\}$ and $\delta(q_f, \lambda) = \cup_{i=1}^{n}\{q_0^1, \ldots, q_0^n\}$.

Check that for any $q \in \cup_{i=1}^{n} Q^i \cup \{q_f\}$, there exists a word w_q such that $\delta(Q_I, w) = \{q\}$. If there exists a word w_0 such that $\delta(Q_I, w_0) = \{q_0\}$ then $\delta(Q_I, w_0\lambda) = \{q_1\}$.

Now, suppose that $\cup_{i=1}^{n} L_i \neq \Sigma^*$ and let $u \in \Sigma^* \setminus \cup_{i=1}^{n} L_i$. Then $\delta(Q_I, u) \cap \cup_{i=1}^{n} Q_T^i = \emptyset$ and therefore, $\delta(Q_I, u\lambda) = \{q_1\}$ and $\delta(Q_I, u\lambda\lambda) = \{q_0\}$.

If $\cup_{i=1}^{n} L_i = \Sigma^*$, for any $u \in \Sigma^*$, $\delta(Q_I, u) \cap \cup_{i=1}^{n} Q_T^i \neq \emptyset$, $\delta(Q_I, u\lambda) = \{q_1, q_f\}$, $\delta(Q_I, u\lambda\Sigma) = \emptyset$ and $\delta(Q_I, u\lambda\lambda) = Q_I$. Therefore, there exists no word w_0 such that $\delta(Q_I, w_0) = \{q_0\}$.

That is, $\cup_{i=1}^{n} L_i \neq \Sigma^*$ if and only if for any $q \in Q_A$, there exists a word $w_q \in \Sigma_A^*$ such that $\delta(Q_I, w_q) = \{q\}$.

Now, associate a new letter y_q to each state $q \in Q_A$ and consider the MA $B = \langle \Sigma_B, Q_B, \iota, \tau, \varphi \rangle$ where

– $\Sigma_B = \Sigma_A \cup \{y_q | q \in Q_A\}$,
– $Q_B = Q_A \cup \{q_b\}$,
– $\iota(q) = 1/(n+1)$ if $q \in Q_I$ and 0 otherwise,
– $\tau(q) = 1$ if $q = q_b$ and 0 otherwise,
– $\varphi(q, x, q') = 1/(\sum_{y \in \Sigma} |\delta(q, y)| + 1)$ if $q, q' \in Q_A$, $x \in \Sigma_A$ and $q' \in \delta(q, x)$,
– $\varphi(q, y_q, q_b) = 1/(\sum_{y \in \Sigma} |\delta(q, y)| + 1)$,
– $\varphi(q, x, q') = 0$ in all other cases.

Check that B is a PA. B is $\mathbb{R}^+$-reduced since for any $q \in Q_A$, $r_{B,q}(y_{q'}) \neq 0$ iff $q = q'$ and $r_{B,q}(\varepsilon) = 0$. B is a PRA if and only if for any $q \in Q_A$, there exists a word $w_q \in \Sigma_A^*$ such that $\delta(Q_I, w_q) = \{q\}$.

Putting all together, we see that an algorithm which decides whether B is a PRA could be used to decide whether $\cup_{i=1}^{n} L_i \neq \Sigma^*$.

As the problem is clearly PSPACE, it is PSPACE-complete. □

It has been shown in [DLT02] that for any polynomial $p(\cdot)$, there exists an NFA $A = \langle \Sigma_A, Q, Q_I, Q_T, \delta \rangle$ which satisfies the following properties:

– for any state q of A, there exists a word $w \in \Sigma^*$ such that $\delta(Q_I, w) = \{q\}$,
– for any state q of A, all words w which satisfy $\delta(Q_I, w) = \{q\}$ have a length greater than $p(|Q|)$.

These NFA are support of PRA which inherit of this property.

So, reduced PRA form a decidable family which is sufficient to generate $\mathcal{S}_K^{fingen}(\Sigma)$ but the membership problem for this family is not polynomial. We can restrict this family to obtain a polynomially decidable family and still sufficient to generate $\mathcal{S}_K^{fingen}(\Sigma)$.

Let $A = \langle \Sigma, Q, \iota, \tau, \varphi \rangle$ be a PRA. A is *prefixial* if for any $q \in Q$, there exists $w_q \in \Sigma^*$ such that $w_q^{-1} r_A = r_{A,q}$ and such that $\{w_q | q \in Q\}$ is prefixial.

It is polynomially decidable whether an MA is a prefixial PRA.

Let $A = \langle \Sigma, Q, \iota, \tau, \varphi \rangle$ be a PRA, and for any $q \in Q$, let $w_q \in \Sigma^*$ such that $w_q^{-1} r_A = r_{A,q}$. Let $W = \{w_q | q \in Q\}$ and let $\overline{W}$ be the smallest prefixial subset of Σ^* which contains W. Let $B = \langle \Sigma, \overline{W}, \overline{\iota}, \overline{\tau}, \overline{\varphi} \rangle$ be the MA defined by:

– $\overline{\iota(q)} = 1$ if $q = \varepsilon$ and 0 otherwise,
– $\overline{\tau}(w) = w^{-1} r_A(\varepsilon)$,

– $\overline{\varphi}(w, x, wx) = w^{-1}r_A(x\Sigma^*)$ for any $x \in \Sigma$,
– $\overline{\varphi}(w_q, x, w_{q'}) = \varphi(q, x, q')$ if $w_q x \notin W$,
– $\overline{\varphi}(w, x, w') = 0$ in all other cases.

It can be shown that B is a prefixial PRA equivalent to A.

4.4 Probabilistic Deterministic Automata

For any $K \in \{\mathbb{R}, \mathbb{Q}, \mathbb{R}^+, \mathbb{Q}^+\}$, the class $\mathcal{S}_K^{fin}(\Sigma)$ can be described by probabilistic deterministic automata.

Proposition 22. *Let $K \in \{\mathbb{R}, \mathbb{Q}, \mathbb{R}^+, \mathbb{Q}^+\}$ and let $p \in K\langle\langle\Sigma\rangle\rangle$. Then, p is a stochastic language over K which has finitely many residual languages iff there exists a K-probabilistic deterministic automaton A such that $p = r_A$.*

Proof. From Prop 14, we can suppose that $K \in \{\mathbb{R}^+, \mathbb{Q}^+\}$.

– Let $p \in \mathcal{S}_K^{fin}(\Sigma)$ and let $Res(p) = \{w_1^{-1}p, \ldots, w_n^{-1}p\}$. Let A be the MA associated with S as in the proof of Prop 18. As there exists $i \in \{1, \ldots, n\}$ such that $p = w_i^{-1}p$, we can suppose that $\alpha_s = 1$ if $s = w_i^{-1}p$ and 0 otherwise. Let $sw_i^{-1}p$. If $x \notin res(s)$, then $\sum_{w\in\Sigma^*} p(w_i x w) = 0$ and since $K \in \{\mathbb{R}^+, \mathbb{Q}^+\}$, this implies that $p(w_i x w) = 0$ for any word w. Therefore, in this case, it is possible to choose $\alpha_{s,s'}^x = 0$ for any $s' \in Res(p)$. When $x \in res(s)$, there exists $j \in \{1, \ldots, n\}$ such that $x^{-1}s = w_j^{-1}p$. In this case, we can choose $\alpha_{s,s'}^x = 1$ if $s' = w_j^{-1}p$ and 0 otherwise.
  Then, check that A is a PDA which generates p.
– Let $A = \langle\Sigma, Q, \varphi, \iota, \tau\rangle$ be a PDA which generates p and let $Q_I = \{q_0\}$. For any $w \in \Sigma^*$, there eixts only one state $q \in Q$ such that $\varphi(q_0, w, q) \neq 0$. Therefore, $Res(p) \subseteq \{r_{A,q} | q \in Q\}$ and $Res(p)$ is a finite state.

□

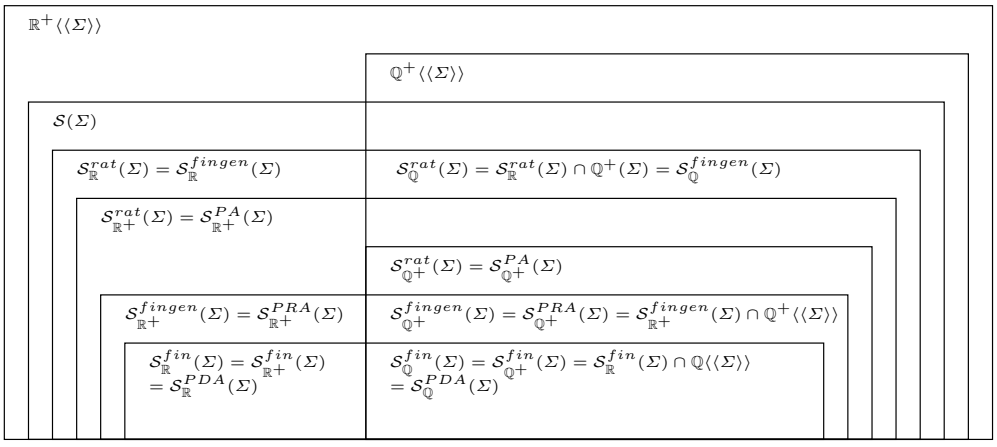

Fig. 8. Inclusion relations between classes of classes of rational stochastic languages.

5 Conclusion

In this paper, we have carried out a systematic study of rational stochastic languages, which are precisely the objects probabilistic grammatical inference deal with. This study, and the results we bring out, whether they are original or derived from former contributions, support our opinion that researches in grammatical inference should be based and rely on formal language theory. Doing this makes it possible to reuse powerful tools and general results for inference purposes. Moreover, this approach may help finding out what particular properties are important for grammatical inference. For example, a learning sample $\{w_1, \ldots, w_n\}$ independently drawn according to a target stochastic language p provides statistical information on the residual languages of p. In order to infer an approximation of p by means of a multiplicity automata A, there should be a structural link between the states of A and the observed data and hence, between the states of A and the residual languages of p. This explains why most results in grammatical inference deal with PDA and PRA, i.e. classes of multiplicity automata for which there exists a strong connection between the states and the residual languages of the stochastic languages they generate. This also explains why there is no useful general inference result about PA: the residual subsemimodule of a rational stochastic language over $\mathbb{R}^+$ or $\mathbb{Q}^+$ may be not finitely generated and hence, no finite set of residual languages can be used to represent it. Moreover, PA admits no natural normal form. On the other hand, the residual subsemimodule of rational stochastic languages over $\mathbb{R}$ or $\mathbb{Q}$ are finitely generated and admit a basis made of residual languages. Even if there exists no recursively enumerable subset of MA capable of generating them, this study has encouraged us to try to find a way to infer these most general stochastic languages. See [DEH06] for preliminary results. We are also currently working on *tree rational stochastic languages*, following a similar approach, in order to deal with tree probabilistic languages inference. This work is still in progress.

References

AW92. N. Abe and M. Warmuth. On the computational complexity of approximating distributions by probabilistic automata. *Machine Learning*, 9:205–260, 1992.

BC03. V. D. Blondel and V. Canterini. Undecidable problems for probabilistic automata of fixed dimension. *Theory of Computing Systems*, 36(3):231–245, 2003.

BR84. J. Berstel and C. Reutenauer. *Les séries rationnelles et leurs langages*. Masson, 1984.

BT00. V. D. Blondel and J. N. Tsitsiklis. A survey of computational complexity results in systems and control. *Automatica*, 36(9):1249–1274, September 2000.

CO94. R.C. Carrasco and J. Oncina. Learning stochastic regular grammars by means of a state merging method. In *ICGI*, pages 139–152, Heidelberg, September 1994. Springer-Verlag.

CO99. R. C. Carrasco and J. Oncina. Learning deterministic regular grammars from stochastic samples in polynomial time. *RAI*, (1):1–20, 1999.

DDE05. P. Dupont, F. Denis, and Y. Esposito. Links between probabilistic automata and hidden markov models: probability distributions, learning models and induction algorithms. *Pattern Recognition: Special Issue on Grammatical Inference Techniques & Applications*, 38/9:1349–1371, 2005.

DE03. F. Denis and Y. Esposito. Residual languages and probabilistic automata. In *30th International Colloquium, ICALP 2003*, number 2719 in LNCS, pages 452–463. SV, 2003.

DE04. F. Denis and Y. Esposito. Learning classes of probabilistic automata. In *COLT 2004*, number 3120 in LNAI, pages 124–139, 2004.

DEH06. F. Denis, Y. Esposito, and A. Habrard. Learning rational stochastic languages. Technical Report ccsd-00019161, HAL, 2006. https://hal.ccsd.cnrs.fr/ccsd-00019161.

dlHT00. C. de la Higuera and F. Thollard. Identification in the limit with probability one of stochastic deterministic finite automata. In *Proceedings of the 5th ICGI*, volume 1891 of *LNAI*, pages 141–156. Springer, 2000.

DLR77. A.P Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society*, 39:1–38, 1977.

DLT02. F. Denis, A. Lemay, and A. Terlutte. Residual Finite State Automata. *Fundamenta Informaticae*, 51(4):339–368, 2002.

DLT04. F. Denis, A. Lemay, and A. Terlutte. Learning regular languages using rfsas. *Theoretical Computer Science*, 2(313):267–294, 2004.

ELDD02. Y. Esposito, A. Lemay, F. Denis, and P. Dupont. Learning probabilistic residual finite state automata. In *ICGI'2002, 6th ICGI*, LNAI. Springer Verlag, 2002.

Fli74. M. Fliess. Matrices de Hankel. *J. Maths. Pures Appl.*, 53:197–222, 1974. + erratum in Vol. 54 (1975).

Gan66. F. R. Gantmacher. *Théorie des matrices, tomes 1 et 2*. Dunod, 1966.

Jac75. G. Jacob. Sur un théorème de Shamir. *Information and control*, 1975.

Sak03. Jacques Sakarovitch. *Éléments de théorie des automates*. Éditions Vuibert, 2003.

Sch61. M. P. Schützenberger. On the definition of a family of automata. *Information and Control*, 4:245–270, 1961.

SS78. Arto Salomaa and M. Soittola. *Automata: Theoretic Aspects of Formal Power Series*. Springer-Verlag, 1978.